

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:34:58 ; Search time 14.65 Seconds
(without alignments)
1188.768 Million cell updates/sec

Title: US-09-905-088A-245

Perfect score: 3732

Sequence: 1 MRLIVAPDLLAWAGATATV.....RKLPSSSGERTLPLPLSQNS 713

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1656.5	44.4	708	4	US-09-131-648-2 Sequence 2, Appl 1
2	442	11.8	1091	3	US-08-986-485-5 Sequence 5, Appl 1
3	408	10.9	1101	3	US-08-986-485-2 Sequence 2, Appl 1
4	407.5	10.9	605	1	US-08-190-802A-49 Sequence 49, Appl 1
5	407.5	10.9	605	4	US-08-477-346-49 Sequence 49, Appl 1
6	407.5	10.9	605	4	US-08-473-088-49 Sequence 49, Appl 1
7	402	10.8	673	4	US-09-063-950-2 Sequence 2, Appl 1
8	397.5	10.7	605	4	US-09-063-950-5 Sequence 5, Appl 1
9	355.5	9.5	1523	4	US-09-182-024A-2 Sequence 5, Appl 1
10	344	9.2	603	1	US-08-190-802A-50 Sequence 50, Appl 1
11	344	9.2	603	4	US-08-477-346-50 Sequence 50, Appl 1
12	344	9.2	603	4	US-08-473-088-50 Sequence 50, Appl 1
13	331	8.9	1525	3	US-09-191-647-2 Sequence 2, Appl 1
14	331	8.9	1525	4	US-09-540-245A-2 Sequence 2, Appl 1
15	331	8.9	1525	4	US-09-540-153-2 Sequence 2, Appl 1
16	324	8.7	1480	3	US-09-191-647-7 Sequence 7, Appl 1
17	324	8.7	1480	4	US-09-540-245A-7 Sequence 7, Appl 1
18	324	8.7	1480	4	US-09-540-153-7 Sequence 7, Appl 1
19	324	8.7	1480	4	US-09-182-024A-5 Sequence 5, Appl 1
20	324	8.7	1480	5	PCT-US91-09055-2 Sequence 2, Appl 1
21	321	8.6	560	3	US-08-592-500-2 Sequence 2, Appl 1
22	321	8.6	560	3	US-08-195-006-2 Sequence 2, Appl 1
23	321	8.6	560	4	US-09-063-950-4 Sequence 4, Appl 1
24	321	8.6	560	5	PCT-US94-07644A-2 Sequence 2, Appl 1
25	316.5	8.5	353	3	US-08-986-485-6 Sequence 6, Appl 1
26	296.5	7.9	649	4	US-09-188-930-305 Sequence 305, Appl 1
27	274	7.3	1112	4	US-09-353-585-2 Sequence 2, Appl 1

28	273	7.3	1112	4	US-09-353-585-3	Sequence 3, Appl 1
29	269	7.2	353	6	5340934-4	Patent No. 5340934
30	263	7.0	1016	4	US-09-180-439-8	Sequence 8, Appl 1
31	261	7.0	282	1	US-08-442-063A-45	Sequence 45, Appl 1
32	261	7.0	307	1	US-08-442-063A-48	Sequence 48, Appl 1
33	261	7.0	333	1	US-08-442-063A-27	Sequence 27, Appl 1
34	261	7.0	342	1	US-08-272-919-2	Sequence 2, Appl 1
35	261	7.0	342	1	US-08-619-916-2	Sequence 2, Appl 1
36	261	7.0	342	5	PCT-US95-08342-2	Sequence 2, Appl 1
37	261	7.0	359	1	US-08-303-238-4	Sequence 4, Appl 1
38	261	7.0	359	1	US-08-458-834-4	Sequence 4, Appl 1
39	260	7.0	968	4	US-09-180-439-3	Sequence 3, Appl 1
40	260	7.0	968	4	US-09-180-439-4	Sequence 4, Appl 1
41	250	6.7	799	4	US-09-180-439-6	Sequence 6, Appl 1
42	245.5	6.6	368	1	US-08-303-238-3	Sequence 3, Appl 1
43	245.5	6.6	368	4	US-08-458-834-3	Sequence 3, Appl 1
44	239.5	6.4	368	6	5340934-2	Patent No. 5340934
45	237.5	6.4	746	5	PCT-US95-10509-2	Sequence 2, Appl 1

ALIGNMENTS

RESULT 1	US-09-131-648-2	Sequence 2, Application US/09131648
Patent No. 6168920	GENERAL INFORMATION:	
APPLICANT: Hallman, Jennifer L.	APPLICANT: Yue, Henry	
APPLICANT: Corley, Neil C.	APPLICANT: Guegler, Karl J.	
APPLICANT: Patterson, Chandra	TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS	
FILE REFERENCE: PF-0576 US	CURRENT APPLICATION NUMBER: US/09/131,648	
CURRENT FILING DATE: 1998-08-10	NUMBER OF SEQ ID NOS: 5	
SOFTWARE: PERL Program	SEQ ID NO 2	
LENGTH: 708	TYPE: PRT	
ORGANISM: Homo sapiens	FEATURE: -	
OTHER INFORMATION: 2687731	US-09-131-648-2	
Query Match	44.4% Score 1656.5; DB 4; Length 708;	
Best Local Similarity	47.2% Pred. No. 2.6e-144;	
Matches 333; Conservative 109; Mismatches 217; Indels 47; Gaps 5;		
17 TATVPVPMHVPCCPOACQIRPMTYPRSSYREATTVCNDLFTLVAPPALPAGTQTLTL 76		Sequence 2, Appl 1
17 TTVQAVDKKVCPCCTCEIRPMTYPRSSYREATTVCNDLFTLVAPPALPAGTQTLTL 76		Sequence 2, Appl 1
77 QNSITRVQSEGLYANLTEDLSQNSFSRDACDFALPQLSLHLEBNQITREDS 136		Sequence 5, Appl 1
77 QNTNIAKIEIS-TDFVNTGLDLSQNNLSVYNIVKMPQLSLYLEENKTEJPERK 135		Sequence 5, Appl 1
137 FAGIASLOETLHNHLYRIAPRAFGLSNLRHLNSNLRALIDSREEMLPNLEITMI 196		Sequence 2, Appl 1
136 LSLSLMOELIYHNHLSITSPAFGLNLRHLNSNLRALIDSREEMLPNLEITMI 195		Sequence 2, Appl 1
197 GGNKVDAILDMNFRPLANDRSIVLAGMNLREISDVLEGLQSLSELSFYDNOLARVPRRA 256		Sequence 4, Appl 1
196 GEMPIRIKDMNFKPLINRSIVIAGINLTETPDNALVGLNLEISISFYDNRLIKVPYVA 255		Sequence 4, Appl 1
257 LRGVETKRLDANKPLQVGGDFRANMLHKEGLANNEELVSDIKFLVNPETLKID 316		Sequence 6, Appl 1
256 LQVYVLFKFLDKNKINIRIGDGSNMLHKEGLINNPEDISIDSLAVDNPIDLRKIE 315		Sequence 6, Appl 1
317 INNPLRSIHPRAFHHLPMQETMLMNNVALSHQQTVESLPNQLQVGLHGNPINCDCV 376		Sequence 2, Appl 1

STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/986,485
FILING DATE: 08-DEC-1997
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/059,448
FILING DATE: 22-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-986-485-2

Query Match 10.9%; Score 408; DB 3; Length 1101;
Best Local Similarity 21.8%; Pred. No. 1.5e+28;
Matches 185; Conservative 124; Mismatches 277; Indels 264; Gaps 30;
8 LILAVY-----AGATATVPVPMWVPCPOCACQIRWYTPRSSYREATTDCN 56
18 LILVLLLELPVYAAAGPRA-----PCAACTCAGDPCTCAGDS-----LDGG 62
57 DFLVAVPALPAGTQTLTLOSNSIVRVDSGLYANTL----- 97
63 GGLALHPEDLPSTWTRSLNKLAEIDPAGFEDLPNLOEVLNNHETLAVASLGGSS 122
98 -----LDLSQNSFSDARDCDFHLPOLLSLHEENOL 129
123 QVVAFLEQQQNRSLDSGLKAYLSLEVLDLNNITTEVRNTYFPHGPIKELNLAGNRI 182
130 TRLEHSTRAGLA-SLOEYLNNHOLYRIAPRAFGSLNLLRLHNSNLLRAIDSRWEM 188
183 GTLEGAFGDGLSRSLTLRLSKNRITQLPVRAFK-LPRLTQLDNNRIRLIEGLTFQGL 241
189 PNLLETIMIGKNVDAIILDMNRPILNRSVLVAGMNLRIISDYALEGLOSLSFYDNO 248
242 NSLEVLKLRNNISKLTGDAFWGLSKMHLHLEYDSLVEVNSGSLYGLTALQLHLNNS 301
249 LARVPR-----ALEQVPGKFLDLNKNPLOORVGPGEFAM 284
302 IARIHKGMSFQKHLHELVSFNNILRLDESLAEISLSVRLSHNSISHAEGAFKGL 361
285 LHLKELGLNNMEELVSID---KFAVLNLPETLKLIDITNNPRLSFIPRAFHLLPOMET 340
362 RSLRYLDLDHNEISGIEDTSGAFSGLEFGH-SKLTLEFGN-KIKSVAKRAFGSLGLEHL 419
341 MLNNALCALHQVIESLPNLOEVLHGNPINCDCVIR-----AANT----- 383
420 NIGGAISVDFADAFVKKNKKEHLISSDSFLCDQLKWLPPMLIGRMLOAFVATCAHP 479
384 -----GTRVRFIEPOSTIC-----AEP----- 400
480 ESKQGOSIFSPPESTFVDDFLKPOITTPPTTMAVGKDIRFTCSAASSSSSPMTFAWK 539

QY 401 -----PDLOR-----LPVREYF-----R 414
DB 540 KQNEVLTNADMENFVHNAQDGEVMEYTIILHROTFBHEGRYQCVITNHEGSTISHA 599
QY 415 EMTDHLPLISPRSPPSLQVAGSESMVLHCRALAEPEPEIYVWPAGLRTPAHAGRY 474
DB 600 RLTVNLVLP-SFTKTPHDITIRTTVARLECAATGHPNPQIAMQKGGTDF-PAABEREM 656
QY 475 RYVP-EGTLELRVTAEEGLITCVANQNLVG--ADKITVYVGRALLDPGDE-----GQ 527
DB 657 HWMPDDVEFITDVKIDDAVYSCTAQNSAGSISANATLTLETSLVPLEDRVVSVE 716
QY 528 GLELRVQETHPYHILLSWTPPNTVSTNLTWSSASSLRGOGATALARPGRTHSVNIT-- 585
DB 717 TVALOCKATG-----NPP-----PRITW-----FKGDRLESL-----TERHHLTPD 752
QY 586 -RL-----QATEYMACLO--VAFADHTQLACVWARTREATSCHRALDRPLAI 634
DB 753 NQILVYQNVVAEDAGRYCTEMSNLTGTERAHSQLSVL-----PAAGC-RKDGTVVGIPTI 806
QY 635 LALAVLLLA 644
DB 807 AVVSTIVTS 816

RESULT 4
US-08-190-802A-49
Sequence 49, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P. O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor binding
INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-190-802A-49
Query Match 10.9%; Score 407.5; DB 1; Length 605;

RESULT 6
US-08-473-089-49
Sequence 49, Application US/08473089
Patent No. 6342368
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Theoreof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSER: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,089
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHEITICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor binding
INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-473-089-49

Query Match 10.9%; Score 407.5; DB 4; Length 605;
Best Local Similarity 25.2%; Pred. No. 6.1e-29;
Matches 152; Conservative 72; Mismatches 173; Indels 205; Gaps 16;

QY 3 LVAALLAMVA-----GATATVPVPMHVPCCPQACQIRPMTTSSST---REATT 52
DB 8 LALALLLSWVALGFRSLGEGADPTGEGEAGPACAC-----SYDDADELS 57
QY 53 VDCNDLFTAVPALPAGTQTLTLLQNSIVRY-----DQSELYG----- 91
DB 58 VFCSSRNITRLPDPVGGTQALMLDGNNTSSVPPAFQWLSLGLNLNLOGQLGLEPQA 117
QY 92 ---LANTLDELISQNSFSDARDCDFHALPOLISLHLENQTLRL----- 133
DB 118 LLGLENLCHLHLERQNSLALGTFAHPPALASLGLSNNRSLRLDGLFEGIGLSMDLNTL 177
QY 134 -----DHSFAGLASIQELLYLNHNOYLRIAPRAFSGLSMLRLHLNSNLLRAIDSRW 184
DB 178 GWSNLAVLPDAFGRGLSRELVLNRLAYLQPALFGLAELRELDLSRNALRAIKANV 237
QY 185 FEMLPNLEIIMGKKNVAI-----LDMN----- 208
DB 238 FVQPLRLKIVLRILIAVAPGAFGLKALRWLDSHNRKAVGLLEDTFPGLLGIRVRL 297
QY 209 -----FRPLANLRSLIYAGMNLREISDYALEGLQSLSELSFYDNLARV----- 252

DB 298 SHNAIASLRPRTEKDLHLEELQLGHNRIQLAERSFEGIGQLEVLTLDHNOQLEVKAGA 357
QY 253 -----PRALEQVPGKFLDKNPQORVGGDANMLHKELEL 292
DB 358 FGLGNVAVMNLGNCRLNLPQVFRGLKLSLHLEGSCSLRPHPTTGTGSLGRPL 417
QY 293 NNMEELVSIDKFAVNLPELTKLIDTNPRSLFIHRAFHLPQWETLMLNNALSAL-- 350
DB 418 KD-NGLVGIEBSQMSIMLLELDLTSN-QLTHLPHRLFOGLKLEYLILSRNRLELPA 475
QY 351 -----HQOVESLPLN-----LQEVGLNG 368
DB 476 DALGPLQRAFMDVSHNRLEALPNSLLAPLGRULSLRNNSLRPTTQPQGLERLMEG 535
QY 369 NPIRDCQVIR---NANAGTRV-RIEIP-----QSTLCAPPDQLRPLPR 409
DB 536 NPWDCGCPKLALRDLALQNPASVPRFVQALCEGDCCOPPAYTYNNITCASPPVGVGLDR 595
QY 410 EV 411
DB 596 DL 597

RESULT 7
US-09-063-950-2
Sequence 2, Application US/09063950C
Patent No. 6225085
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas A.
TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: MEI-019
CURRENT APPLICATION NUMBER: US/09/063,950C
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 673
TYPE: PRT
ORGANISM: Homo sapiens
US-09-063-950-2

Query Match 10.8%; Score 402; DB 4; Length 673;
Best Local Similarity 24.9%; Pred. No. 2.3e-28;
Matches 194; Conservative 73; Mismatches 306; Indels 206; Gaps 24;

QY 4 LVAPLLAMVAGATATVPVPMHVPCCPQACQIRPMTTSSSYREATTVDCNDLFLAV 63
DB 7 LLLPLLLLAG-----PVGCCPSGCGCS-----QPTVFTTARQCTTV 46
QY 64 PPALPAGTQTLTLLQNSIVRYDQSELYGTYLANLTLELDSQNSFSDARDCDFHALPOLLSLH 123
DB 47 PRDVPDPDFVGLVFNNGTMDAGSFAGLPGLQLDLSQNOQLASLPSGVFQPLANLSND 106
QY 124 LEENQTLREHDSFAGLASIQELLYLNHNOYLRIAPRAFSGLSMLRLHLNSNLLRAIDSR 183
DB 107 LTAHLHEITMETFERGLERLERYLGKNRIRHIGGARDPTDLRLLEKLQDNEELALPPL 166
QY 184 WEMLPNLEIIMGKNKYDATIDMNFRLANLRSLIYAGMNLREISDYALBELQSLSELS 243
DB 167 ---RLPRLLLDLSHNSILA-LEPGITDANEALRLAGLGQQLDEGLFSRLRLHDLD 222
QY 244 FYDNLARVPRALRQVGLKFLDKNPQORVGGDFANMLHKELELNNMEELVSIDK 303
DB 223 VSDNQLERVP-----PVIR----- 236
QY 304 FALVLPBELTKLIDTNPRSLFIHRAFHLPQWETLMLNNALSALHQOVESLPLNQE 363
DB 237 ---GLRGITRLRLAGNRRIRQLRREDLAGLAALQELDVNSISLQALGDDSGLEPRLRL 292
QY 364 VGLHGNPIRCQCVIRMAN--ATGTFVRFTIEPOSTLCAEPPLDQLRLVAREVFRREM----- 416

QY 237 --OSLESISFYDNOQLARPRALRQVPGKLELDLNKPNLQRVGPGFANMLHKEGL-- 292
DB 305 LPEGIYEIRLEONSISKAIPAGAFYTKRKIDISKNOISDIAPDAFQGLKSLSTLYLG 364
QY 293 NNNEELVSDIKFALVNLPELTIKDITNNPRLSFTHPRAFHHLPOMETLMNNNALSLAHQ 352
DB 365 NKITEIAKGLFDELVS-----QILLNANKINCLRVNTFDDLDLNLSTLYDKLOTISK 420
QY 353 QYVESLPNIQEVGHGPNIRCDVIRMANATGVRFRFIEBOSTLCAEPDLOQLPVREV- 411
DB 421 GLFAPLOSIOITLHONPFCDCDKLWL-ADYLQDNPITERSGARCSSPRRLANKRISQIK 479
QY 412 --PER-----EMTDHC-LPLISPR-----SFPSSLOY 435
DB 480 SKKFRSGSSEYDRSRSSSECFMDLCPKRCRCGTIVDCSNOKLVRIPIHLPYVTDLRL 539
QY 436 ASGESVYLHCRALAEPEPEIYWTTPA-----GLRLTRAHAG 471
DB 540 NDNEVSVLATGTGFKLPNIRKINLSNNKIKVEGAFDGAASVQELMLTGNLETVH-G 598
QY 472 RRYR-VYPEGTLELRVTAEAGLYTCAQNLV-----GADTKT 509
DB 599 RYFRGLSGKTKMLR-----SNLISCVSNDFPAGLSVRLSLYXNRITTTTTPGAFITL 652
QY 510 VSVYVGRALLQGRDRGQGLELRVQETHPHYHILLSVTPPTVSTN 555
DB 653 VSPVHHKPPVOP-----LOLOL-----PLAWLGKWLKRRIYSGN 687

RESULT 10

US-08-190-802a-50
Sequence 50, Application US/08190802a

Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190,802a

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 603 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Insulin-like growth factor bind.

INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33

US-08-190-802a-50

Query Match 9.2%; Score 344; DB 1; Length 603;

Best Local Similarity 23.3%; Pred. No. 4, 5e-23;

Matches 141; Conservative 63; Mismatches 194; Indels 208; Gaps 16;

QY 8 LLLAWVA-----GATATPVVPMHVPCPCACQIRMYVPRSSYREATTVDGN 56
DB 13 LLAFWALRPGCHQGDPPGASAAE-----GPOCPVACTS-----HDDYDELVSFCS 61
QY 57 DFLVAPPALPAGTQTLLQNSIVRV-----DOSELGY 91
DB 62 SKNLTLPPDIPYSTRALMDGNLSIPSAARQNLSDFLMLQGSWLRSLRQALLG- 120
QY 92 LANLTELDSQNSFSARCDFFALPQLSLHLEENQTLRF----- 133
DB 121 LQNLVYLHERNRLRMVAGLFTHPPSLASTLSMILGRLEGLRQGLSHLMDLNGWN 180
QY 134 -----DHSFAGLASQELLYNHNOLYRIAPRFGSLNLRLELNSLLRAIDSRPFEM 187
DB 181 SLVLPDVTYFQGLGNLHELVLGKNKLTLYQPALFCGLGELREILSRNALRSYKANVEVH 240
QY 188 LPNLETIMIGNKVDAL-----LDMN----- 208
DB 241 LPRLOKTYLDNMLITVAPGAFGLMKALRWLDLSHNRVAGLMEDTPGGLGLHYLRALHN 300
QY 209 -----FRPLANRLSVLAGMNLREISDYALBGLQSLSELSFYDNOIARV----- 252
DB 301 AIALSRPRTFKDLHLEELQGLGNHRIQGERFEGLEGQLEVLTLDNDQITVEYVAGAFSG 360
QY 253 -----PRALQVPGKLELDLNKPNLQRVGPGFANMLHKEGLNMM 295
DB 361 LFNVAVMNLSGNCRLSPRPFVQGLDKLSLHLEHSCIGHVRLHFFAGLSGRRLFLRD- 419
QY 296 EELVSDIKFALVNLPELTIKDITNNPRLSFTHPRAFHHLPOMETLMNNNALSL----- 350
DB 420 NSISIEBQSLAGLSLELDDLTN-RLTHLRLQRLFOGIGHLEYLLISTNOULTLSAEVL 478
QY 351 -----HQ-----QYVESLPNIQEVGHGPNIR 372
DB 479 GPLORAFMIDISHNLETLAEGLSLGRVYLSLRNNSLOFFSPQGLERLMDLNDPMD 538
QY 373 CDCVIR-----WANATGTVRRIEP-----QSTICARPDQRLPVREPR 414
DB 539 CSCPLALRDFALQNPVGVRPFVQVCEGDCQPYTYNNITCASPANVSGLDLVDSET 598
QY 415 EMTDHC 420
DB 599 HPV-HC 603

RESULT 11

US-08-477-346-50
Sequence 50, Application US/08477346

Patent No. 6262023

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:34:58 ; Search time 35.32 Seconds

(without alignments)
2242.231 Million cell updates/sec

Title: US-09-905-088a-245

Perfect score: 3732

Sequence: 1 MRLIVAPLLAWAGATATV.....RKLPSSGEETILPPLSQNS 713

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq.032802.*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3732	100.0	713	20	AAV13385
2	3732	100.0	713	22	AAAB80253
3	1993	53.4	377	22	AAAB92765
4	1695	45.4	716	21	AAAB3472
5	1695	45.4	716	22	AAU12427
6	1695	45.4	716	22	AAAB27234
7	1695	45.4	716	22	AAAG50965
8	1695	45.4	716	22	AAAG67530
9	1672	44.8	716	22	AAAM40376
10	1656.5	44.4	705	21	AAAB42513
11	1656.5	44.4	705	22	AAAM78823

12	1656.5	44.4	708	20	AAV13355	Amino acid sequenc
13	1656.5	44.4	708	21	AAV70030	Human extracellular
14	1656.5	44.4	708	22	AAAM39309	Human polypeptide
15	1656.5	44.4	708	22	AAU12320	Human PRO220 polyp
16	1656.5	44.4	708	22	AAAB80223	Human PRO220 prote
17	1654.5	44.3	719	22	AAV79807	Human protein SEQ
18	1654.5	44.3	719	22	AAAM4095	Human polypeptide
19	1640	43.9	707	20	AAV02379	Polypeptide identi
20	1504	40.3	292	22	AAE03524	Human secreted pro
21	875.5	23.5	431	22	AAAB94521	Human protein sequ
22	755.5	20.2	273	22	AAAB93523	Human protein sequ
23	546.5	14.6	592	22	AAE09437	Human sbgTrango79a
24	492.5	13.2	620	22	AAAB74705	Human membrane ass
25	490.5	13.1	614	20	AAAB84596	Amino acid sequenc
26	489.5	13.1	620	20	AAV13357	Amino acid sequenc
27	489.5	13.1	620	22	AAU12333	Human PRO227 polyp
28	489.5	13.1	620	22	AAAB80225	Human PRO227 prote
29	468.5	12.6	606	22	AAE06799	Human neuronal gut
30	467.5	12.5	579	22	AAE06804	Mature human neuro
31	442	11.8	1091	18	AAAM41641	Sequence used in d
32	442	11.8	1091	20	AAV08099	Murine glial cell
33	442	11.8	1091	20	AAV08010	Mouse LiG-1 protei
34	442	11.8	1091	21	AAV97833	Murine LiG-1 prote
35	438.5	11.7	640	20	AAV08100	Human PRO331 prote
36	438.5	11.7	640	20	AAAB85722	Novel protein (Cio
37	438.5	11.7	640	20	AAV13394	Amino acid sequenc
38	438.5	11.7	640	21	AAAB24407	Human PRO331 prote
39	438.5	11.7	640	21	AAV70673	Human PRO331 prote
40	438.5	11.7	640	22	AAU12355	Human PRO331 polyp
41	438.5	11.7	640	22	AAU00826	Human immune respo
42	438.5	11.7	640	22	AAAB80262	Human PRO331 prote
43	438.5	11.7	640	22	AAAB65292	Human PRO331 prote
44	438.5	11.7	640	22	AAAB53089	Human angiotensin
45	432.5	11.6	540	22	AAAB65612	Drosophila melanog

ALIGNMENTS

RESULT 1	AAV13385	standard; Protein; 713 AA.
ID	AAV13385;	
XX	AAV13385;	
XX		
DT	25-JUN-1999	(first entry)
XX		
DE	Amino acid sequence of protein PRO293.	
KW	Secreted protein; transmembrane protein; human; enterocolitis;	
KW	Zollinger-Ellison syndrome; gastrointestinal ulceration;	
KW	congenital microvillus atrophy; skin disease; cell growth;	
KW	abnormal keratinocyte differentiation; porriasis; epithelial cancer;	
KW	Parkinson's disease; Alzheimer's disease; ALS; neuropathy;	
KW	fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;	
KW	anti-thrombotic; wound healing; tissue repair.	
OS	Homo sapiens.	
XX		
PN	WO9914328-A2.	
XX		
PD	25-MAR-1999.	
XX		
PF	16-SEP-1998;	98MO-US13330.
XX		
PR	25-NOV-1997;	97US-0066840.
XX		
PR	17-SEP-1997;	97US-0059113.
XX		
PR	17-SEP-1997;	97US-0059115.
XX		
PR	17-SEP-1997;	97US-0059117.
XX		
PR	17-SEP-1997;	97US-0059119.
XX		
PR	17-SEP-1997;	97US-0059121.
XX		
PR	17-SEP-1997;	97US-0059122.
XX		
PR	17-SEP-1997;	97US-0059184.

KW antiarthritic; antifertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation.
XX
OS Homo sapiens.
PN MO200104311-AL.
XX
PD 18-JAN-2001.
XX
PF 22-FEB-2000; 2000MO-US04414.
XX
PR 07-JUL-1999; 99US-0143048.
PR 26-JUL-1999; 99US-0143698.
PR 28-JUL-1999; 99US-0146222.
PR 08-SEP-1999; 99MO-US20594.
PR 13-SEP-1999; 99MO-US20944.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
PR 05-OCT-1999; 99MO-US23089.
PR 29-NOV-1999; 99MO-US28214.
PR 30-NOV-1999; 99MO-US28313.
PR 16-DEC-1999; 99MO-US30095.
PR 20-DEC-1999; 99MO-US30911.
PR 20-DEC-1999; 99MO-US30999.
PR 05-JAN-2000; 99MO-US00219.
XX
XX (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvarski E, Fong S, Gao W, Geiber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kijavini JF;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI.
XX
DR WPI: 2001-081051/09.
DR N-PsDB: AAF72414.
XX
PT Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease) -
XX
XX
PS Claim 1; Fig 86; 393pp; English.
XX
CC The present sequence is one of sixty one novel secreted and
CC transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemias such as coronary
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC diabetes and retinal disorders such as retinitis pigmentosa.
CC The PRO nucleic acids have applications in molecular biology, including
CC use as hybridization probes, and in chromosome and gene mapping.
XX
SQ Sequence 713 AA:

Query Match 100.0%; Score 3732; DB 22; Length 713;
Best Local Similarity 100.0%; Pred. No. 1e-290;
Matches 713; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLVAPLLAVVAGTATVPVPMVHPCPCACQIRPWYPRSSRYRATVYDCNDL 60
DB 1 MLLVAPLLAVVAGTATVPVPMVHPCPCACQIRPWYPRSSRYRATVYDCNDL 60
QY 61 TAVPALPAGTGTLLQSNISIVVDQSELGYLANLELDSNPSFSDARDCFHALPOL 120
DB 61 TAVPALPAGTGTLLQSNISIVVDQSELGYLANLELDSNPSFSDARDCFHALPOL 120

QY 121 SLHLEENQTRLEDHSPFAGLASIQEELYLNHNOYRIAPRAFSGSLNLRHLNSLRAT 180
DB 121 SLHLEENQTRLEDHSPFAGLASIQEELYLNHNOYRIAPRAFSGSLNLRHLNSLRAT 180
QY 181 DSRFEMLPNLEIIMIGKVFDAIILDMNRPPLANLSVLAGMRLREISYALEGLOSLE 240
DB 181 DSRFEMLPNLEIIMIGKVFDAIILDMNRPPLANLSVLAGMRLREISYALEGLOSLE 240
QY 241 SLSPYDNOQLARVRRALEQVPGKFLDLNKNPQLQVPGSPFANLHLKEGLNMEELYS 300
DB 241 SLSPYDNOQLARVRRALEQVPGKFLDLNKNPQLQVPGSPFANLHLKEGLNMEELYS 300
QY 301 IDKFAVLNLEPRLKIDITNPNRSLFIHPRAFHLPQMETLNNNALSAALHOOTVESLPN 360
DB 301 IDKFAVLNLEPRLKIDITNPNRSLFIHPRAFHLPQMETLNNNALSAALHOOTVESLPN 360
QY 361 LQEVGLHGNPRLCDVYRMANATGTRVRETEPOSTLCAEPDLCRLPVRVPREMTDHC 420
DB 361 LQEVGLHGNPRLCDVYRMANATGTRVRETEPOSTLCAEPDLCRLPVRVPREMTDHC 420
QY 421 LPLISPRSPPSLOVAGSESMVHCRALAPPEPEITVWTPAGRLRPAHAGRRYRVEG 480
DB 421 LPLISPRSPPSLOVAGSESMVHCRALAPPEPEITVWTPAGRLRPAHAGRRYRVEG 480
QY 481 TLRLRYTAEAGLYTCVAONLVGADTKTVSVYVGRALDPGDEGGLELRVOETHPYH 540
DB 481 TLRLRYTAEAGLYTCVAONLVGADTKTVSVYVGRALDPGDEGGLELRVOETHPYH 540
QY 541 ILLSWTPPTVSTNLTWSSASSLRGOGATPALRLPGTHSYNITRLLQATEWYACLOVA 600
DB 541 ILLSWTPPTVSTNLTWSSASSLRGOGATPALRLPGTHSYNITRLLQATEWYACLOVA 600
QY 601 FADAHQOLACVMARTEATGCHRALDRGLAIILAVILLAAGLAHAGTGOPKRGVC 660
DB 601 FADAHQOLACVMARTEATGCHRALDRGLAIILAVILLAAGLAHAGTGOPKRGVC 660
QY 661 GRRLPPAMAFMGMSAPSVRVASAPLVLPNPGKRLPRSSSEGETLPLPSSONS 713
DB 661 GRRLPPAMAFMGMSAPSVRVASAPLVLPNPGKRLPRSSSEGETLPLPSSONS 713

RESULT 3
AAB92765
ID AAB92765 standard; Protein; 377 AA.
XX
AC AAB92765;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:11234.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
OS Homo sapiens.
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.

DR MPI: 2000-572371/53.
 XX N-PSDB: AAC58637.
 PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
 PS Claim 33; Fig 11b; 309pp; English.
 CX
 CC The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
 CC and peripheral nervous systems, demyelinating diseases of the central
 CC and bowel disease, gluten-sensitive enteropathy and Whipple's disease,
 CC autoimmune or immune-mediated skin diseases, allergic diseases,
 CC immunological diseases of the lung, and transplantation associated
 CC diseases including graft rejection and graft-versus-host-disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.
 CC XX
 SQ Sequence 716 AA:
 Query Match 45.4%; Score 1695; DB 21; Length 716;
 Best Local Similarity 53.6%; Pred. No. 3.5e-127;
 Matches 335; Conservative 104; Mismatches 166; Indels 20; Gaps 5
 29 CPGCACGACGRRPYTSSRSRYEATYDCNCLFLTAVRPALPAGTQTLLDSSNSTR-VDOS 87
 32 cpqlccetprrfpcqstyreatvdcndrlrltripsnsdqtqyllgsnmlakvtd-- 89
 88 ELGYLANLTLELDLSONSFSDARDCEFHAPOLLSLTLENOGLTRLEDHSPAGIASQELY 147
 90 elqqlfnlteldlfsqgnfnlkevglaqltltlhleeqtlemdcdyqlqdslnqely 149
 148 LNHNOGLYRIAPAFSGLSNULRLHNSLRLAIRDSRFEMRLPNTLEITMGKNKVDAILDM 207
 150 lnhogstlsatshafaglknlrlrhnsnklkvidsrfdstpnlelimgenprvgilnm 209
 208 NFRPLANLRSLVLAAGNLRLEISDYALEGQSSESJSFYNOCLARVRPRALREOVPGSLFTD 267
 210 nfrplannlrslvlagmlyltdlprnalaivgldsleslsgfynklkvqpalqkynulkfd 269
 268 LNKNPLOARPGPDFANMLHKLKELGANNMEELYSIDKFAIVNLPRLTKLIDITNNPRLSFTH 327
 270 lnknpnlkigedfknnmlrllkeilgimmgelvsdvdiadalpeltklleaetnmpklsyth 329
 328 PRAFNHLRQMETLMLNNALSAHQQTVESELPNLQEVGLHGNPIRDCCYIIRANATGRV 387
 330 prafnhlrmetlmlnnalnaahqqtveselpnlqevglhgnpirdccyihnanatgrv 389
 388 RFIEQSTFLCAERPOLQRLPVNEVFRREKTDHCLPLISRSFSPSLQVANGSGEWLHCRA 447
 390 rfieqstflmcfmppekyghqvkevilqdsseqclpmisidnsfpnrlndvigtvfldfca 449
 448 LAERPELTYWYTPALRLYLRTPAHAGRRYRVRPGSTELRNVTPAEELGGLYVCVQNLVGATP 507
 450 laerpeeltywtpalrllylrtpahagrryrvrpgstelrnvtpaeelggllyvcvqnlvgatp 509
 508 KTVSVVVGRRALLQPRDGGQGLELKVQETNHRPHILLSWTPTPNPTSTNLSSAS--SLRG 566
 510 kvtvsvvvgrrallqprdgqglelkvqetnhrphilswtptpnptstnlkwsasacmkldh 565
 567 OGATPLADLRPGSTHSYNTIRLLQATLEYNACIQAVAFADAHHTOLACYAWARTKEA----- 618

Db	566	phlytlarypvdheynlthlpgstyecltvsnlngtqkscvnyttknaafavdsd	625
Oy	619	----TSCRALGDRPGILAILALAV 639	
Db	626	getstalaawgsmfavislaslav 650	
RESULT	5		
AAU12427			
ID	AAU12427	standard; Protein; 716 AA.	
XX	AC	AAU12427;	
XX	DT	24-OCT-2001 (first entry)	
XX	DE	Human PRO1338 polypeptide sequence.	
XX	KW	Human secretory and transmembrane; PRO; mammalian; cancer; lung;	
XX	KW	breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;	
XX	KW	cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;	
XX	OS	adipocyte; A-peptide; factor VIIA; gene therapy.	
XX	OS	Homo sapiens.	
XX	PN	WO200140466-A2.	
XX	PD	07-JUN-2001.	
XX	PF	01-DEC-2000; 2000MO-US32678.	
XX	PR	01-DEC-1999; 99MO-US28301.	
XX	PR	01-DEC-1999; 99MO-US28634.	
XX	PR	02-DEC-1999; 99MO-US28551.	
XX	PR	02-DEC-1999; 99MO-US28564.	
XX	PR	02-DEC-1999; 99MO-US28565.	
XX	PR	09-DEC-1999; 99US-0170262.	
XX	PR	16-DEC-1999; 99MO-US30095.	
XX	PR	20-DEC-1999; 99MO-US30911.	
XX	PR	30-DEC-1999; 99MO-US30999.	
XX	PR	06-JAN-2000; 2000MO-US31243.	
XX	PR	06-JAN-2000; 2000MO-US00277.	
XX	PR	11-FEB-2000; 2000MO-US00376.	
XX	PR	18-FEB-2000; 2000MO-US03365.	
XX	PR	18-FEB-2000; 2000MO-US04341.	
XX	PR	22-FEB-2000; 2000MO-US04342.	
XX	PR	22-FEB-2000; 2000MO-US04414.	
XX	PR	24-FEB-2000; 2000MO-US04914.	
XX	PR	24-FEB-2000; 2000MO-US05004.	
XX	PR	01-MAR-2000; 2000MO-US05061.	
XX	PR	20-MAR-2000; 2000MO-US07377.	
XX	PR	21-MAR-2000; 2000MO-US07532.	
XX	PR	30-MAR-2000; 2000MO-US08439.	
XX	PR	17-MAY-2000; 2000MO-US13705.	
XX	PR	22-MAY-2000; 2000MO-US14042.	
XX	PR	30-MAY-2000; 2000MO-US14941.	
XX	PR	02-JUN-2000; 2000MO-US15264.	
XX	PR	10-NOV-2000; 2000MO-US30873.	
XX	PA	(GETH.) GENENTECH INC.	
XX	PI	Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;	
XX	PI	Gertsen ME, Goddard A, Godowski PJ, Gunney AL, Sherwood S;	
XX	PI	Smith V, Stewart TA, Tunas D, Watanabe CK, Wood WI, Zhang Z;	
XX	DR	WPI; 2001-408281/43.	
XX	DR	N-PSDB; AAS21499.	
XX	PT	Isolated, secretory and transmembrane PRO polypeptide used to detect	
XX	PT	other PRO polypeptides, link bioactive molecules to cells expressing	
XX	PT	PRO polypeptides, and detect the presence of mammalian tumours e.g.	
XX	XX	lung, breast, prostate, cervical -	

```

Db      210 nfkprlanrslvlagmymttdiprgnalygladstestiydnklvkvpgtalqkvpnlkfld 265
Oy      268 LNKRNPLQVRGPDFAANMLHLKEIGLNNNEELVSDIKFALVNLPELTKFKDITNNRPLSFII 327
Db      270 lnkorpikhgqdfkmmrlrklkeldginnmgelvsvdryaldnlpeltklateatmpklsyih 329
Oy      328 PRAFNHPLQMETLMLNNNALSALHQQTVESLPLNLEWGLHGNPRLRCDOVIRMANATGTRV 387
Db      330 rlatfsvpaleslmmalnnaiaiykctvesipnlreitsinsnprrcdcvlhwlnsoktnl 389
Oy      388 REIEQSTLCAERPDQQLPVREVPFRKMTDHCPLISPNSEFPRLDVASGESNWLHCRA 447
Db      390 rimeplnftcmampreykqhgykewllqdseseqclpmashdsfprlnvdiqtvtvfidcra 449
Oy      448 LAERPELTYWTPRGLRLTTPRHAGRRRYRVYDEGTELELRMTAEAGLGYTCVAQNLVADOT 507
Db      450 maepereplywtrpignkltvetellcdkxkylasegelltsnldqtedsgyrttcvaqnvqgd 509
Oy      508 KTVSVVVGRRALLQGRDEGQLELRVQDETNYHRIILSMWYPRPNTVSNLTMSSAS-SLRG 566
Db      510 rvaalkvngtll-----dgtyglkkyvngqtshesllvswkxnsnwmvsnlkwssatmkldn 565
Oy      567 QGATRLARLPFGTYSYNTIRLLQATETWACIQAFAADAHQLOACVMARKEA----- 618
Db      566 phltcyarvdyvheynulthipqstdyevcltvsnlhgqkcevnvntknaafavdisd 625
Oy      619 -----TSCRHALGDRPGLTALITALLAV 639
Db      626 getataalaavmgsmfavlastasiav 650

```

RESULT	7	
AAB50965		
ID	AAB50965	standard; Protein; 716 AA.
XX		
AC	AAB50965;	
XX		
DT	21-MAR-2001	(first entry)
XX		
DE	Human PRO1338 proteln.	
XX		
KW	Human; PRO; cyostatic; nootropic; neuroprotective; respiratory general	
KV	antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;	
KW	PRO agonist; cancer; inflammatory disorder; Immunological disorder.	
OS	Homo sapiens.	
PN	WO200073348-A2.	
XX		
PD	07-DEC-2000.	
XX		
PF	30-MAY-2000;	2000WO-US14941.
XX		
PR	02-JUN-1999;	99WO-US12252.
PR	22-JUN-1999;	99US-0140650.
PR	23-JUN-1999;	99US-0141037.
PR	20-JUL-1999;	99US-0144758.
PR	01-SEP-1999;	99WO-US20111.
PR	08-SEP-1999;	99WO-US20594.
PR	29-OCT-1999;	99US-0162506.
PR	30-NOV-1999;	99WO-US28313.
PR	01-DEC-1999;	99WO-US28634.
PR	02-DEC-1999;	99WO-US28551.
PR	16-DEC-1999;	99WO-US30095.
PR	20-DEC-1999;	99WO-US30999.
PR	06-JAN-2000;	2000WO-US00376.
PR	11-FEB-2000;	2000WO-US03565.
PR	18-FEB-2000;	2000WO-US04341.
PR	18-FEB-2000;	2000WO-US04342.
PR	02-MAR-2000;	2000WO-US05841.
PR	03-MAR-2000;	2000US-0187202.
PR	10-MAR-2000;	2000WO-US06319.

PR	15-MAR-2000; 2000OWO-US06884.
PR	30-MAR-2000; 2000OWO-US08439.
PR	17-MAY-2000; 2000OWO-US13705.
XX	
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC,
PI	Shelton DJ, Smith V, Watanabe CK, Wood WI;
XX	
DR	WPI: 2001-016509/02.
DR	N-PSDB: AHC91567.
XX	
PT	Twenty eight nucleic acids encoding PRO polypeptides which are useful
PT	for treating various tumors, e.g. breast cancer, and other
PT	inflammatory, angiogenic and immunological disorders -
XX	
PS	Claim 31; Fig 30; 188pp; English.
XX	
CC	The present sequence is one of twenty eight novel PRO polypeptides. The
CC	PRO polypeptides and their agonists, including antibodies, peptides, and
CC	small molecule agonists, may be used to treat various tumours, e.g.,
CC	cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
CC	cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
CC	central nervous system cancer, melanoma or leukaemia. They are also
CC	useful for treating other disorders such as neuronal, glial, astrocytal,
CC	hypothalamic and other glandular, macrophagal, epithelial, stromal and
CC	biostoreic disorders, and inflammatory, angiogenic and immunological
CC	disorders.
XX	
XX	
Sequence	716 AA;
50	

Query Match	45.4%	Score 1695;	DB 22;	Length 716;
Best Local Similarity	53.6%	Pred. No. 3.5e-127;		
Matches 335;	Conservative 104;	Mismatches 166;	Indels 20;	Gaps 5
QY	CPPOCAQIRPYPTRRSSYREATVDCNDLFTAVPPALPACTOFTLLQSNSTIVR	VDQS	87	
Db	32	cpqlcvceitprwftprgstcyreatvtdcndlrtripsnlsotqyllqismakvtd--	89	
QY	88	ELGYANTLEIDLSONSFSDARDCFHALLPOLLSLHLEENQTLRUEHDSFAGLAEOLY	147	
Db	90	elqgfntleldsfsgnfntkewglanjtctihleengftemtqdyldqslshqely	149	
QY	148	LNHNQLYIAARAFSGSLNLRHLHNSULNRIDSRWEMLEPNLITLMIGKKVDALIDM	207	
Db	150	lnhnqslstshahaglkmltrllrhlnshnskikwdsrtdwspnlellmagenprvgildm	209	
QY	208	NEPRPLANRSLVLAGMNRREISDYALDELQSQSESESPFDNOLARPRRLAEQVPGKFLD	267	
Db	210	nfrplanrslrvlagmnyltdipgnalvgldstseslfsfndkivkprqalqayvpnlkfld	269	
QY	268	LNKNPDLQVPGDFPANNMLHLKELGLNNBELVSDKFLVNLPELTKLDITNNPRLSFTH	327	
Db	270	lnknprlnkigqdfxkmrlrkelfglnmgelvsvqrdyldnlpeltkileatnprklsyln	329	
QY	328	PRAFHLPOMETLMLNNNALSLAHQOATYESLPNLOEVLHNGPRLTCDQVIRAKNNTGFRV	387	
Db	330	rlafstwpaleslmmalnalaikyktlveslprnlreislshnpirdcvclhmlnsknknl	389	
QY	388	RFIEPOSTLCAEPDLORLPVREYFPRMTDHCPLRISPSFPELQVAGSGSMVLCRA	447	
Db	390	rflmepslmfcampcykghykevlldqssesqclpmishdsfprnlndvldgtvldicta	449	
QY	448	LAEPPELTYWTPRGLRLTPRNAGRRKYVYRGGTIELRRVTAEEAGLYTCVAQNILVADT	507	
Db	450	maepepeiywtpfrygnklyvetclsdkyklseagflesnldsgrytcvaqnvqgadt	509	
QY	508	KTVSVVVGARALQCGRGEGGELRVQETHRYHLLSVTPPNRYSTLMTSSAS-SLRG	566	
Db	510	rvaclikvngtll-----dgtvklkyvqchesllsvskvsnvmtcslrkwsaslmkldn	565	
QY	567	QGATATLARLPGRHSHYNTRLQATEWACQVAFADNHTQLACVWARTKEA-----	618	

```
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
PA
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang Y;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI, 2001-442253/47.
DR N-PSDB; AAI59532.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries.
XX
PS Example 6; SEQ ID NO 3521; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nucleotide,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Prager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 716 AA;
XX
Query Match 44.8%; Score 1672; DB 22; Length 716;
Best Local Similarity 53.1%; Pred. No. 2.5e-125;
Matches 332; Conservative 105; Mismatches 168; Indels 20; Gaps 5;
OY 29 CPOCAGQIRPWYTPRSSREATTVCNDPLFLAVPPALPAGQOTLLQSNSTVR-VDOS 87
DB 32 cpqlccwcelrpfptgqstyreattvdcnolrlrrpsnsdsdqylllsnlnaktvd-- 89
OY 88 ELGYANTLTEDLSQNSFSFSDARDCDFHALPOLSLHLERNOUTLRDHSFAGIASQERY 147
DB 90 elqqlfnlteldfsqnfnfkveqjlanltqtlthleengqitemdyclqdslnqely 149
OY 148 LNHNLQYRAPAFSGLSNLRLHLNSNLRAIDSFWFEMLPNLETLMTGKNKDAIDLM 207
DB 150 lnhnqlstisafteglxnlrlhlnsnklkyidsrftwdsfpyldlilmganpvgldlm 209
OY 208 NRPPLANTLSLVLAGKNLRREISDYALEGLESSESYFNOLARVRRALAEVPGKFLD 267
DB 210 nrlpplantlsylvlagmyltldipgnalvgidseslstfynklvxpqalqgvnplkfld 269
OY 268 LKNKPLQRYVGPDFANMLHLKELGNNMEELVSIDKFAVLNLPETLKLDITNNPRLSFTH 327
DB 270 lknkplrkqegdfkmrlrklkelglnmgelvsvdryaldnlpelklkaatnpxkysyh 329
OY 328 PRAFHHLQMEFLMLNNNAASLHQQTYVASLPNLOEVGLHGNPFRCDVIRMANATGTRV 387
DB 330 prafhhlqmeflmlnnnaaslhqqtyvaslpnloevglhgnpfrcdvirmantgrv 387
```

```
DB 330 rlafrsvpaleslnlnnaalnalyqktvesipnlreishnsplrcdcvlnhrlnsktni 389
OY 388 RFRFPOSTLCAEPDLDRLPYRREVPFREMNDHCPLISPSFPLSLQVASESNVHLHORA 447
DB 390 rfrfpostlcaepdldrlpyrrevpfremndhcxplispsfplslqvasesnvhlhora 449
OY 448 LAEPEPEIYWTTPAGRLRTPAHAGRRRYRVYDEGFLREIRYTAERAGLYTCVANOVLGADT 507
DB 450 maepepeiywttpagrlrtpahagrrryrvydegflreirytajeraglytcvavngadt 509
OY 508 KTVSVVVGRRALQPGRDEGQELRVOETHPYHLLSMWTPPNTVSTNLMSSAS-SLRG 566
DB 510 rlatikvngtll----dgtvlklyvqteshslvswkxnsnmwntnlnkwsatlmkldn 565
OY 567 QGATLALRLPRGTHSVITPFLQATEYACLOVAFAADAHQOLACWARTREA----- 618
DB 566 phlytarvpvdherylthlqpsdyevcltvsnlhngqkscvntknaafavdsd 625
OY 619 ----TSCRRALGDRPGLIATLALAV 639
DB 626 getstalaavmgsmfavlsiaslav 650
RESULT 10
AAB42513
ID AAB42513 standard; Protein; 705 AA.
AC AAB42513;
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2277 polypeptide sequence SEQ ID NO:4554.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerrary; antiproliferative; antiparkinsonian; nocrotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
KW Homo sapiens.
XX
OS WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000MO-US08621.
XX
PR 31-MAR-1999; 990US-0127607.
PR 02-APR-1999; 990US-0127636.
PR 05-APR-1999; 990US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
PI WPI, 2000-602362/57.
DR N-PSDB; AAC76722.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 3752-3754; 5507bp; English.
```

QY	17	TATVTVVWVHNPCCPQCAQIRPWTVTPRSSSRREATVTDVCDNLETTAVVPALPACTQTLTL	76		
Db	14	tlvavavkxvdcprlccctceirvfwtpcrisymeastvdcnclgllltfparipantqllll	73		
QY	77	QSNSTVRYVDQSEFLGVLANTVELDLSQNSFSDARQCDFFALPQQLSLHNEQNLRLRDEHS	136		
Db	74	qtnnlakleys tdrpvnltqldslsqmllsvtrlnvkkmpqlllsvyleenkkelpck	132		
QY	137	FAGLSAQELVLTNNQYLRIPARAFSGLSNLLRLHLSNLSNLRADISWFEKLPMLFIMI	196		
Db	133	ltselnlqelylnhnlletspgefignhllrlhlnsnrlqmnkskfwclapnllelmi	192		
QY	197	GGNKVDALLDNKFRPLARLSVLVLAGMKLRISDYLAEGLQSELSFYDQNLARVRA	256		
Db	193	genpflrlrkdmfkpllnrlrlvlaglnlrlpdaalvglenleslsfydrlllkvrhva	252		
QY	257	LEQVGLKFLDLNKKPRLQVGPGRANLHLKELGINNMEELVSDKFAVLNLPETKLD	316		
Db	253	lqkvnlrlfldlnknprrlrrtfgdfrsmhlhkeigimmpelsldslavdnlpdrkile	312		
QY	317	ITNNRLSFIRPRAFNHLPQMETLMLNNALSAHQOVESTLPMLOEVLGHGPIRCDCV	376		
Db	313	atnprlslglnpnafrlrpkleslmnsalsalghgteslprlkelslshpnrdcv	372		
QY	377	IRMAAATRTVRFTEPOSTLCAEPPDQLQRLVKEVPRREMTDHLPLISPSPPSLQVA	436		
Db	373	lrmmnmktnlrflnepdslfcvdppefiggvnrvghrlrmdelcplrlapdesfpsnlve	432		
QY	437	SGESVWLHCRALAEPERETVWVTAGLRFLPAHAGRGVRVYRPGETELRLRVTAEEGLYT	496		
Db	433	agsyvsfcrstaerqpelylrlpsgqklpnrlctdkfyhsegcltdingvtrpeglyt	492		
QY	497	CVAQMLVADRTKTVSVVVGRLALOPGRDESGGELFRVQETHPHAILLSWTPTNVTSTNL	556		
Db	493	ciatcllygadsklsvmkvxdgsifpq--dnngslhikrlqdgansvltvswkaskxlksav	549		
QY	557	TWSSASSLRGGAGALALRLPGCTHSYNTLRLQATEWACLOQVAFADANHQLQACWAARTK	616		
Db	550	kvtafvtkenshaagqsarlpdvkvvynltnhpsteykcidiprltyqnrkkcnvntk	609		
QY	617	-----EATSCHRALGDRPGULALALAVLLLAAGLAANLSTGQPRKGVGGR-	663		
Db	610	glhpdqkeyeknnttlmacrlglllglyiclscl-----spemncdghs	657		
QY	664	-----PLPRAAFMGWSAPSVKRVVSAPLVLPWN	691		
Db	658	yvrnylqkprtfalagelylppllnlweagkxstskvkatvlgjptn	703		
RESULT 12					
ID	AA013355	AA013355 standard; Protein: 708 AA.			
XX	AA013355;				
XX	AA013355;				
DT	25-JUN-1999	(first entry)			
XX					
DE	Amino acid sequence of protein PRO220.				
XX					
KM	Secreted protein; transmembrane protein; human; enterocolitis;				
KM	Zollinger-Ellison syndrome; gastrointestinal ulceration;				
KM	congenital microvillus atrophy; skin disease; cell growth;				
KM	abnormal keratinocyte differentiation; psoriasis; epithelial cancer;				
KM	Parkinson's disease; Alzheimer's disease; ALS; neuropathy;				
KM	fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;				
KM	anti-Chromobiotic; wound healing; tissue repair.				
XX					
OS	Homo sapiens.				
XX					
PN	W09914328-A2.				
XX					
DD	25-MAR-1999.				

XX	16-SEP-1998;	98MO-US19330.
PF		
XX	25-NOV-1997;	97US-0066840.
XX	17-SEP-1997;	97US-0059113.
PR	17-SEP-1997;	97US-0059115.
PR	17-SEP-1997;	97US-0059117.
PR	17-SEP-1997;	97US-0059119.
PR	17-SEP-1997;	97US-0059121.
PR	17-SEP-1997;	97US-0059122.
PR	18-SEP-1997;	97US-0059184.
PR	18-SEP-1997;	97US-0059263.
PR	18-SEP-1997;	97US-0059266.
PR	15-OCT-1997;	97US-006125.
PR	17-OCT-1997;	97US-006285.
PR	17-OCT-1997;	97US-0062287.
PR	21-OCT-1997;	97US-0063486.
PR	24-OCT-1997;	97US-0062814.
PR	24-OCT-1997;	97US-0062816.
PR	24-OCT-1997;	97US-0063045.
PR	24-OCT-1997;	97US-0063120.
PR	24-OCT-1997;	97US-0063121.
PR	24-OCT-1997;	97US-0063127.
PR	24-OCT-1997;	97US-0063128.
PR	27-OCT-1997;	97US-0063329.
PR	27-OCT-1997;	97US-0063327.
PR	28-OCT-1997;	97US-0063541.
PR	28-OCT-1997;	97US-0063542.
PR	28-OCT-1997;	97US-0063544.
PR	28-OCT-1997;	97US-0063549.
PR	28-OCT-1997;	97US-0063550.
PR	28-OCT-1997;	97US-0063564.
PR	29-OCT-1997;	97US-0063435.
PR	29-OCT-1997;	97US-0063704.
PR	29-OCT-1997;	97US-0063732.
PR	29-OCT-1997;	97US-0063738.
PR	29-OCT-1997;	97US-0063738.
PR	29-OCT-1997;	97US-0063734.
PR	29-OCT-1997;	97US-0064215.
PR	29-OCT-1997;	97US-0063735.
PR	31-OCT-1997;	97US-0063870.
PR	31-OCT-1997;	97US-0064103.
PR	03-NOV-1997;	97US-0064809.
PR	07-NOV-1997;	97US-0064809.
PR	12-NOV-1997;	97US-0065186.
PR	17-NOV-1997;	97US-0065846.
PR	18-NOV-1997;	97US-0065693.
PR	21-NOV-1997;	97US-0066120.
PR	21-NOV-1997;	97US-0066364.
PR	24-NOV-1997;	97US-0066772.
PR	24-NOV-1997;	97US-0066466.
PR	24-NOV-1997;	97US-0066770.
PR	24-NOV-1997;	97US-0066511.
PR	24-NOV-1997;	97US-0066453.
XX		
PA	(GETH) GENENTECH INC.	
XX		
PI	Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;	
XX	WPI: 1999-229533/19.	
DR	N-PSDB; AAX52226.	
DR		
XX		
PT	New isolated human genes and polypeptides used in, e.g. treatment of	
PT	gastrointestinal ulceration	
XX		
PS	Claim 12; Fig 26; 320pp; English.	
XX		
CC	AAV13344-403 represent secreted and transmembrane human proteins.	
CC	The cDNA sequences are obtained from cDNA libraries, prepared from	
CC	fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.	
CC	The encoded polypeptides have specific uses based on their homology to	
CC	known polypeptides, e.g. PRO211 and PRO217 can be used for disorders	
CC	associated with the preservation and maintenance of gastrointestinal	
CC	mucosa and the repair of acute and chronic mucosal lesions	

C;Comment: This factor is structurally related to proinsulin and have insuline-like meta

A:Gene: CESF121D12.9a
A:Map position: 4
A:Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2

Db 178 -----PVOA-----FRSLALQAMTLANKIHADIYACGNISLV 213

QY 241 SLSEYDNOIARVRRALAEQVPGKFLDLNKPLORVGPGDFANMLHKEGLNMEELVS 300

Db 214 VLHNNRHISLKGKCFDGLHSETLDLNNNDEF-PTAIKRLSNKEGLGPHS-NNIRS 271

QY 301 IDKFAVYVNPBELTKLITNNPRLSFTHPRAFHHLPOMETLMLNNA----- 346

Db 272 IPRRAVGNPSLTIHFYDNP-IOFVGSAFOHLPRLTLTLNGASHITEPMLTGTATL 330

QY 347 -----LSALHOOVESLPNLOEGLHGN 369

Db 331 ESUTLTGAKISSLPQAVCDLPLMLQVLDLSYN 362

RESULT 8

T42218

slit-1 protein homolog - rat

N:Alternate names: MEGR4 protein

C:Species: Rattus norvegicus (Norway rat)

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000

C:Accession: T42218

R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.

Genomics 51, 27-34, 1998

A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs

A:Reference number: Z14126; MUID:98360089

A:Accession: T42218

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1531 <NAK>

A:Cross-references: EMBL:AB011530; NID:g3449289; PIDN:BAA32460.1; PID:g3449290

A:Experimental source: strain Sprague-Dawley; brain

C:Genetics:

A:Gene: MEGR4

Query Match

Best Local Similarity 9.7%; Score 360.5; DB 2; Length 1531;

Matches 136; Conservative 76; Mismatches 216; Indels 149; Gaps 17;

QY 8 LLL---AWAGATATATPVVPMWHPVPCPOCACQIRPWYTPRSSYREATTVDGNDLFLTAVP 64

Db 20 LLLMAAMRLGATA-----CPALCTC-----TCTVDCHEGTGLQAIIP 56

QY 65 PALPAGTQTLLONSIYRVDSLOGLYLANTELDSQNSFSPARDCEFHALLPOLLSLH 124

Db 57 KIPRTERLELNGNNITRIHKNDFAGLKQLVQLQLEMENIGAVEGADDKKELERLAL 116

QY 125 EENQITREDSFAGLASLOELYLNHNOLYRIAPRAFSGLSNLRHLNLSNLRATDSRM 184

Db 117 NKNQLOVYPELLFQNNQALSRDLSENLSQAVPRKAFRGATDLKQLDKNOISCIIEGA 176

QY 185 FEMLNLETLIMIGKNKVAIILDMNFRPLANLRSL----- 218

Db 177 FRALGLELVTLLNNNNITITPVSSFNHMPKLTFRHLSNHLFCDCHLANLSQWLRORPIT 236

QY 219 -----VLGNNLREI--SDYALBGLQSLSEISFYDNOIARVPR----- 254

Db 237 GLFTQCGSPASIRGLNVAEYQKSEFSCS-----QGAAYVPACTLSSGSCPAMC 286

QY 255 -----RALEQVPG-----LKFIDLNNKPLQRVGPGDFANMLHKEGLNMEEL 298

Db 287 SCNSGIVDCRGKGLTAIPANLPETWTEIRLELNGIKSIIPGAFSPYRKLRRIIDLGN-NQI 345

QY 229 VSIQKFAVNPBELTKL-----DTNNNR-----LSFIHPRAFPHILP 335

Db 346 AEIAPDAQGLKSLNSLVLYGNKITDLRGVYGGLYTQLQLLNANKINCIPDAPFODLO 405

QY 336 QMETLMLNNALSLAQOOTESLPNLOEGLHGNPIRCDVCYIRMANAGTRARFIEPOST 395

Db 406 NLSLSTLDNKTQSLAKTFTSLRAIQTLHLAQNPFICDCKNLKWL-ADPLKRNPLETTGA 464

QY 396 LCAEPPDQLRLPVREVPREM-----TDHCLPLLSRSPSPSLQVASESVLHCRALAE 450

Db 465 RCASPRRLANKRIGQIKSKFRCSAKQOYFIP--GTEDYHLSNCTSDVACPHRCRC--- 519

QY 451 PEPETIYVTPAGLRLT--PAHAGRRVYDEGLELR 485

Db 520 ---EASVVECSGLKLSKIPER-----IPQSTTELR 546

RESULT 9

JE0176

orphan G protein-coupled receptor precursor - human

C:Species: Homo sapiens (man)

C>Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000

C:Accession: JE0176

R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.

Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A:Title: Identification and cloning of an orphan G protein-coupled receptor of the GJ

A:Reference number: JE0176; MUID:98308104

A:Accession: JE0176

A:Molecule type: mRNA

A:Residues: 1-907 <MCND>

A:Cross-references: GB:AF062006; NID:g3366801; PIDN:AAC28019.1; PID:g3366802

C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.

C:Genetics:

A:Gene: HG38

A:Map position: 12q22-23

F:1-21/Domain: signal sequence #status predicted <SIG>

F:562-583/Domain: transmembrane #status predicted <TM1>

F:594-616/Domain: transmembrane #status predicted <TM2>

F:639-660/Domain: transmembrane #status predicted <TM3>

F:681-701/Domain: transmembrane #status predicted <TM4>

F:725-744/Domain: transmembrane #status predicted <TM5>

F:768-791/Domain: transmembrane #status predicted <TM6>

F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match

Best Local Similarity 9.6%; Score 360; DB 2; Length 907;

Matches 134; Conservative 70; Mismatches 189; Indels 96; Gaps 14;

QY 3 LLAVALLLAMVAGATATPVVPMWHPVPCPOCACQIRPWYTPRSSYREATTVDGNDLFLTAVP 62

Db 9 LLSLPVLLQATGSSSPRGVLLR--GCPTRHCHCE-----PDGRMLLRVDCSDGLSE 59

QY 63 VPPALPAGTQTLLONSIYRVDSLOGLYLANTELDSQNSFSPARDCEFHALLPOLLSL 122

Db 60 LPSNLSVFTSYLDLSNNNISQILLPNPLPSLRLLEELRLAGNALTYIPKGAFTGLYSIKVYL 119

QY 123 HLEENQITREDSFAGLASLOELYLNHNOLYRIAPRAFSGLSNLRHLNLSNLRATDSRM 182

Db 120 MLONNQLRHVPPEALONLSLOSIRLDANHISVPPSCSGLSLRLMLWDNALTEI-- 177

QY 183 RWFEMLPNLEIIMIGKNKVAIILDMNFRPLANLRSLVLAGMNLREISDYALBGLQSLSEI 242

Db 178 -----PVOA-----FRSLALQAMTLANKIHADIYACGNISLV 213

QY 243 SLSEYDNOIARVRRALAEQVPGKFLDLNKPLORVGPGDFANMLHKEGLNMEELVSID 302

Db 216 VLHNNRHISLKGKCFDGLHSETLDLNNNDEF-PTAIRLSNKEGLGPHS-NNIRSIP 273

QY 303 IDKFAVYVNPBELTKLITNNPRLSFTHPRAFHHLPOMETLMLNNA----- 346

Db 274 IPRRAVGNPSLTIHFYDNP-IOFVGSAFOHLPRLTLTLNGASOITERPDLTGTANLES 332

QY 347 -----LSALHOOOTESLPNLOEGLHGNPIRCDVCYIRMANAGTRARFIE--POSTIC 397

Db 333 LFTLGAQSSLPQYTCNQGLPNLQVLDLSYN-----LLDDPSFSVC 373

QY 398 AEPDQLRLPVREVPREM-TDHCPLPLSPRSPSPSLQVASESVLHCRALAEPEPEY 456

Db 374 OK---LQKIDLRHNEIYEIKVDTFOOLLSLR-----SLNLAAMKXIAITH-----PNMF 418

QY 457 WTPPAGLR 465

R:Botelclair, Y.R.; Seto, D.; Hsieh, S.; Hurst, K.R.; Ooi, G.T.
Proc. Natl. Acad. Sci. U.S.A. 93, 10028-10033, 1996
A:Title: Organization and Chromosomal Localization of the gene encoding the mouse acid 1
A:Reference number: J06128; MUID:96413591
A:Accession: J06128
A:Molecule type: DNA
A:Residues: 1-603 <BOU>
A:Cross-references: GB:U06900; NID:91621612; PID:AB1720.1; PID:91621613
C:Comment: This protein is a serum protein and it is of the ternary complex in the phys
C:Genetics:
A:Gene: als
A:Map position: 17

Query Match 9.2%; Score 342.5; DB 2; Length 603;
Best Local Similarity 22.0%; Pred. No. 1.3e-15;
Matches 139; Conservative 69; Mismatches 190; Indels 233; Gaps 15;

QY 8 LLLAWA-----GATATVPVPMHVPCCPOCACQIRWYPRSSYREATTVDGN 56
DB 13 LIAFWALGPCYLQGTDPASADAE---GPCPVTCTCSY-----DDYDELSEVCS 61
QY 57 DEFTAVPPALPAGTQTLTLLQSNIVR-----DQSELCG 91
DB 62 SRNLTQLPGICPSTRALMDGNNLSIPSAFQNTSLDEFLNLQGSWLRSLERPOALLG- 120
QY 92 LANTELDSQNSFSDARCDFFALPOLSLHLENQTLRE----- 133
DB 121 LQNTYLHLERNLRLSLAAGLFPHHPSTSLSGNNLLRLRLEGLRGLSHLMDLNGN 180
QY 134 -----DHFAGLASIQEYLNNQLYRAPRFSGLSMLRLHLSNLRADISWFEM 187
DB 181 SLVLPDVTVOGIGNHLYLGNKLTLYLPALCGIGELREHLSRNALRSVKANVFTH 240
QY 188 LFNLETIMGKNVDI-----LDNN----- 208
DB 241 LPRLOKLYDRNLITRVAPRALGKALKRDLDSHNKRVAGLLEDFPGLLGHVLRAN 300
QY 209 -----FRPLANLRSVLAGMNLREISDYALEGSLSESYFYDNLARV----- 252
DB 301 ATTSIRPRFKDLHLEELQGHNRRLQGEKTFEGLQLEVLTLNDNQHIEYKVAFG 360
QY 253 -----PRALEQVPGKFLDLNKNPLQVGGDFANMLHKEKELANN 295
DB 361 LFNVAVMNLGCLRSLEPHVFOGLRHLHSHCLGRIRLHPFAGLSGLRFLRD- 419
QY 296 EELVSLDKFALVNLPELTQDITNNPRLSFIRPRAFHILPQMETMLNNALSALHQ--- 352
DB 420 NSISSTIEGSLAGSELLELDTAN-QTLHLPRLQFQIGOLEYLLSNNOQLTMSEDVL 478
QY 353 -----QTVESLPNLOEYGLHGNPIR 372
DB 479 GPLQAFNIDLSHNLETPAEGLFSSGLRKLRLNNSLQTFVPOGELERMLDANPD 538
QY 373 CDCVIR-----WANATGFRFRIEP-----OSTCAEPDQLRPLVEVPR 414
DB 539 CSCPKALRDEALQMPGVYRFVQTVCEGDDCQPYTYTNNITCAGPANSGLDLDI- 595
QY 415 EMTDHLPLISRPSPPLSIQVAGSGSMVLHC 445
DB 596 -----SETLFFVHC 603

RESULT 13
B3665
slit protein 2 precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 19-May-2000
C:Accession: B3665
R:Robberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2169-2187, 1990
A:Title: Slit: an extracellular protein necessary for development of midline glia and co
A:Reference number: A3665; MUID:91099665

A:Accession: B3665
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1469 <ROR>
A:Cross-references: GB:X53959
C:Genetics:
A:Gene: FlyBase:sl
A:Cross-references: FlyBase:FBgn0003425
C:Superfamily: unassigned EGF-related proteins; EGF homology; leucine-rich alpha-2-g
F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>
F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>
F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>
F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>
F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>
F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>
F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>
F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
F:1028-1061/Domain: EGF homology <EGF2>
F:1068-1099/Domain: EGF homology <EGF2>
F:1115-1148/Domain: EGF homology <EGF1>

Query Match 8.7%; Score 324; DB 2; Length 1469;
Best Local Similarity 24.7%; Pred. No. 9e-14;
Matches 103; Conservative 67; Mismatches 169; Indels 78; Gaps 11;

QY 29 CPPCACQIRPMYTPRSSYREATTVDNDLFFAVPPALPAGTQTLTLLQSNIVRQDSE 88
DB 519 CPAMCHC-----EGTVDCGRRUKETPRDPLHTFELLNDNEGRISDGG 565
QY 89 LQYLANTELDSQNSFSDARCDFFALPOLSLHLENQTLREDSFAGLSQELYL 148
DB 566 L-----FGRPLHVKLELKRQNLGIEPNAEGASHIDQLD 602
QY 149 NNQLYRIAPRAFSGLSNLRHLNSNLRADISWFEMLPULETIMGKNVDAILDN 208
DB 603 GENKIETSKNMFELGHQKTLNLYDNOISCVPSFEHLNLSLIMLASNPENC--NCH 660
QY 209 FRPLAN-LRSLVLAGMNL-----EISDVALEGLSLE----- 240
DB 661 LAMFECYAKKLSGAAACGAPSKVYQVQIDLPHESEKSSENSEGCLGDGYCPSPCT 720
QY 241 -----SLSEFYDNLARVPRRALEQVPG-LKFLDLNKNPLQVGGDFANMLHKEKELANN 295
DB 721 CTGTVAAASRNQLEIPR-----GIPATSELYLENEETLQHYERIRHLRSLTRIDLSN- 775
QY 296 EELVSLDKFALVNLPELTQDITNNPRLSFIRPRAFHILPQMETMLNNALSALHQQT 355
DB 776 NQITLISVTFANLTKLSTLITSYN-KIQCLQRLHALSGLNLRVYSLGNRISMPEGSF 834
QY 356 ESLPNOEYGLHGNPIRDCVIRMANGTGRTFEEPOSTICAEPPDQLRPLVEVPR 412
DB 835 EDLKSITHALGSLNPLVCDGLKMF-SDWIKLDVPEGIARCAEPQKDKIILSTP 890

RESULT 14
A3665

C:Comment: The amino end of the intact protein is blocked.
 C:Comment: This protein is absent in Bernard-Soulier syndrome.
 C:Genetics:
 A:Gene: GDB:GPI5
 A:Cross-references: GDB:230236; OMIM:173511
 A:Map position: 5pter-5qter
 C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
 C:Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane prot

Query Match 8.6%; Score 321; DB 2; Length 560;
 Best local similarity 26.6%; Pred No. 3.5e-14;
 Matches 141; Conservative 57; Mismatches 227; Indels 106; Gaps 16;

```

QY 9 LIAWAGATATVAVVPMVHPCPCACQIRPWYTPRSSYREATTVDCNDLF-LTAVPPAL 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 LICAVIGLIRAGP-----PCCPPACKC-----VFRDAACSGGDVARISAL--GL 48

QY 68 PAG-TOTLLQSNSTYRVQSEIGYLANLTEDLSONSFSDARDCCDFHALPQLLSHLEE 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49 PNLTHILLFGMGVGLQSGFSG-MTVLQRLMISDSHISAVAGTFESDLIKIKTLRLSR 107

QY 127 NOTLRDEHSPAGIASIOELYNHNLRYRARAEGSLNLRHLNSNLRRAIDSRMFE 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 NKITHLPGLLDKMYLLEQLEFDHNLRGIDQMFQKLVNLELALNQDLPLPASLFT 167

QY 187 MLENLEILMIGNKVDAIIDMNERPLANRSLVLAGMNLREISDYALEGQSLESFYD 246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 NLENKLLDLSGNLTHLPKGLGAQAKLERLLHSNRLVSLDGLNSIGALTLEQFHR 227

QY 247 NOLARVPRALAEQVPGKFLDLNK-----NPLQRVGPGDFA 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 NHRSIAPGAFDRPLPMLSSITLSRNHLAFPSALFLHSNLTLLTFENPLAELPGVLEF 287

QY 283 NMLHKEELGNNEELVSDKFAIVNLPETKIDITNNRSLFIHRAFHLPQMETML 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 EMGGLOELMLNR-TQLRTLPAAPFRNLSRLRYLGVTLSPRLSALPGAFQGLGELQVLA 346

QY 343 NNNALSAL-----HOQ-----T 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 HSNGLTALPDGLRGIGLRLROYSLRRNRRLRALFRNLSSLESYQLDHNQLETLPQDV 406

QY 355 VESLPNLOEVGLHGNPIKDC---VIRMANATGTRVREIEPOSTLCAEPPDLQRLPYRE 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 FGALPRLTEVLIGHNSMRDCGIGPFLGWLRLGLVGGEEPPR--CAGPGAHAGLPLMA 464

QY 411 VFPREMTDHCPLDISPSPPLQVASGESMTLHCALAEPEPEIT-WVTP 460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 465 LPEGDA--EC---PGRGPPPRPADSSSEAPVH-PALAPNSSEPPVWMAOP 509
  
```

Search completed: August 31, 2002, 14:36:40
 Job time: 102 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:35:28 ; Search time 13.6 Seconds
(without alignments)
2029.929 Million cell updates/sec

Title: US-09-905-088a-245
3732

Sequence: 1 MRLVAPLLLAIVAGATATV.....RKLPRSSGERTLLPPLSQNS 713

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3718	99.6	713	1	GAC1_HUMAN
2	407.5	10.9	605	1	AUS_HUMAN
3	397.5	10.7	605	1	AUS_PAPHA
4	344	9.2	603	1	AUS_RAT
5	342.5	9.2	603	1	ALS_MOUSE
6	340.5	9.1	567	1	GPV_MOUSE
7	324	8.7	1480	1	SLIT_DROME
8	321	8.6	560	1	GPV_HUMAN
9	314	8.4	905	1	TLR3_MOUSE
10	313.5	8.4	567	1	GPV_MOUSE
11	298.5	8.0	361	1	CHAD_BOVIN
12	298	8.0	536	1	CBP8_HUMAN
13	287	7.7	904	1	TLR3_HUMAN
14	280.5	7.5	662	1	GAP8_HUMAN
15	279	7.4	782	1	CHAO_TRICA
16	276.5	7.4	357	1	PGS2_CHICK
17	272	7.3	1315	1	CHAO_DROME
18	270	7.2	360	1	PGS2_PIG
19	270	7.2	1050	1	TLR7_MOUSE
20	269.5	7.2	1032	1	TLR9_MOUSE
21	268.5	7.2	1032	1	TLR9_HUMAN
22	268	7.2	354	1	PGS2_MOUSE
23	268	7.2	360	1	PGS2_BOVIN
24	266	7.1	1097	1	TOLL_DROME
25	261	7.0	359	1	PGS2_HUMAN
26	261	7.0	360	1	PGS2_HUMAN
27	260	7.0	977	1	PGS2_RABIT
28	259	6.9	331	1	PLIB_AGBXL
29	258	6.9	354	1	PGS2_RAT
30	253.5	6.8	1049	1	TLR7_HUMAN
31	253	6.8	360	1	PGS2_CANFA
32	253	6.8	372	1	PGS1_HORSE
33	252	6.8	1032	1	TLR8_MOUSE

34	251.5	6.7	369	1	PGS1_SHEEP	046390	ovis aries
35	251	6.7	360	1	PGS2_HORSE	046542	equus caball
36	250.5	6.7	369	1	PGS1_BOVIN	P21809	bos taurus
37	250.5	6.7	682	1	CONN_DROME	001819	drosophila
38	249	6.7	369	1	PGS1_MOUSE	P28653	mus musculus
39	248.5	6.7	369	1	PGS1_MOUSE	002678	mus musculus
40	248.5	6.6	369	1	PGS1_MOUSE	P47853	rattus norv
41	246.5	6.6	368	1	PGS1_HUMAN	P21810	homo sapien
42	246.5	6.6	2026	1	CYAA_YEAST	P08678	saccharomyc
43	242.5	6.5	582	1	SHO2_HUMAN	Q9UG13	homo sapien
44	242.5	6.5	582	1	SHO2_MOUSE	088520	mus musculus
45	241	6.5	342	1	LDM_BOVIN	Q05443	bos taurus

ALIGNMENTS

RESULT	ID	STANDARD	PRT	713 AA
1	GAC1_HUMAN			
AC	075325;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Glioma amplified on chromosome 1 protein precursor.			
GN	GAC1			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Glial tumor;			
RX	MEDLINE=98324709; PubMed=9662332;			
RA	Malloy B., Almeida A., Zhu X.X., Vogt N., Tyagi R., Muller M.,			
RA	Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;			
RT	"GAC1, a new member of the leucine-rich repeat superfamily on			
RT	chromosome band 1q32.1, is amplified and overexpressed in malignant			
RT	gliomas."			
RL	Oncogene 16:2997-3002(1998).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).			
CC	-1- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.			
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: AF030435; AAC39792.1; -			
DR	MIM: 605492; -			
DR	InterPro: IPR003006; IG_MHC.			
DR	InterPro: IPR003598; IG_C2.			
DR	InterPro: IPR001611; LRR.			
DR	InterPro: IPR000483; LRR_Nterm.			
DR	InterPro: IPR000372; LRR_Cterm.			
DR	InterPro: IPR003592; LRR_out.			
DR	InterPro: IPR003591; LRR_typ.			
DR	Pfam: PF00047; Ig_1.			
DR	Pfam: PF00560; LRR; 10.			
DR	Pfam: PF01463; LRRCT; 1.			
DR	Pfam: PF01462; LRRNT; 1.			
DR	PRINTS: PR00019; LRRICHPT.			
DR	SMART: SM00408; IGC2; 1.			
DR	SMART: SM00370; LRR; 6.			
DR	SMART: SM00082; LRRCT; 1.			
DR	SMART: SM00013; LRRNT; 1.			
DR	SMART: SM00369; LRR_Typ; 2.			

DR SMART; SM00013; LRRNT; 1.
 KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 605 INSULIN-LIKE GROWTH FACTOR BINDING
 FT REPEAT 53 73 LRR 1.
 FT REPEAT 74 96 LRR 2.
 FT REPEAT 98 120 LRR 3.
 FT REPEAT 121 144 LRR 4.
 FT REPEAT 145 168 LRR 5.
 FT REPEAT 170 192 LRR 6.
 FT REPEAT 193 216 LRR 7.
 FT REPEAT 217 240 LRR 8.
 FT REPEAT 242 264 LRR 9.
 FT REPEAT 265 288 LRR 10.
 FT REPEAT 289 312 LRR 11.
 FT REPEAT 313 336 LRR 12.
 FT REPEAT 337 360 LRR 13.
 FT REPEAT 361 384 LRR 14.
 FT REPEAT 386 408 LRR 15.
 FT REPEAT 409 432 LRR 16.
 FT REPEAT 433 456 LRR 17.
 FT REPEAT 458 480 LRR 18.
 FT REPEAT 482 504 LRR 19.
 FT REPEAT 505 530 LRR 20.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 605 AA; 66034 MW; F6562A23CBE918F6 C664;

Query Match 10.9%; Score 407.5; DB 1; Length 605;
 Best Local Similarity 25.2%; Pred. No. 6; Se-22;
 Matches 152; Conservative 72; Mismatches 173; Indels 205; Gaps 16;

QY 3 LLVAPLLAMVA-----GATATVYVYVWVHPVPCQACQIRPWTYPRSSY---REATT 52
 DB 8 LALALLLSVALGPRSLGADPGTGEAGPACRACAC-----STDDADELS 57
 QY 53 VDCNLEFLTAVPALPAGTQTLLOSNSIVRY-----DQSELGY----- 91
 DB 58 VECSSRNLTFLPDGVGFGQALMDGNMNSVPPAFAQNLSSGFNLGQGLGLEPQA 117
 QY 92 ---LANTFELDSQSFSDARCDGFHALPOLSLHEENQTLRL----- 133
 DB 118 LLGLENLCHLHERNOLNSLAGTFAHTPAPALASLGNSNRLSLRLEGLFEGLSIMDLNL 177
 QY 134 -----DHSFAGIASLOELLYLNHNOLYRIAPRAFSGLSNLRHLNLSNLRATDSRW 184
 DB 178 GWNSLAVLPDAAFRGISRELVLGNRLAYQPALFSGLALRELDLSRNALRAIKAVY 237
 QY 185 FEMAPNLEIIMGGKKVDAI-----LDMN----- 208
 DB 238 FVQLPRIQKLYLDRMLIAAVAGAFGLKALRWLDLSHRVAGLLEDFPGLGLRYLRL 297
 QY 209 -----FRPLANLRSLVLAGMNLRESDVALBSGLSLESLFYDNLARY----- 252
 DB 298 SHNAIASLRPTFFKDLHLLEELQLSHNRIROLAERSFEGLOLEVLTLDHNOLOEKVKA 357
 QY 253 -----PRRALEQVPGKLEFLDKNPNLORVPGDFANMLHLKEGL 292
 DB 358 FLGLTNVAVMNLGNCANLBPQVFRGKGLSHLSLBSSCLRIKPRHTGSGLRRLFL 417
 QY 293 NNMEELVSIKFAVLNDELTKLDTNNPRLSFTHPRAFHHLPMQETMLNNAALSAI-- 350
 DB 418 KD-NGIVGIEEQSLMGSLAELELDLTSN-QTLHPRLRFQOGKGLKLEYLLSSNRRLAELPA 475
 QY 351 -----HOQTVESLPN-----LQEVGLHG 368

DB 476 DALGPIQRAFWLDVSHNRLEALPNLSLLAPLGRRLYLSTRNNSLRTFPPGGLERLMEG 535
 QY 369 NPTRCOCVIR-----WANTATRY-RFIEP-----QSTLCAPPDQLRLPVR 409
 DB 536 NPWDCCGPELALRDFALQNSAVPRVQAICEGDCCQPPATYNNITCASPPVEVGLDLR 595
 QY 410 EV 411
 DB 596 DL 597
 RESULT 3
 ID ALS_PAPHA STANDARD; PRT; 605 AA.
 AC 002833;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 01-MAR-2002 (rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
 DE IGFALS OR ALS.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 OX NCBI_TaxID=9557;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RC MEDLINE=97040714; PubMed=8886027;
 RA Delhanty P., Baxter R.C.;
 RT "The cloning and expression of the baboon acid-labile subunit of the
 RT insulin-like growth factor binding protein complex.";
 RT Biochem. Biophys. Res. Commun. 227:897-902(1996).
 CC -I- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
 CC IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
 CC -I- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
 CC IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: Extracellular.
 CC -I- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL, S83462; -; NOT_ANNOTATED_CDS.
 DR HSSP; P23945; LXUN.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR: 19
 DR Pfam; PF01463; LRRCT; 1.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PR00019; LEORICHRPT.
 DR SMART; SM00370; LRR: 2.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_typ; 11.
 KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 605
 FT REPEAT 53 73 LRR 1.
 FT REPEAT 74 96 LRR 2.
 FT REPEAT 98 120 LRR 3.
 FT REPEAT 121 144 LRR 4.
 FT REPEAT 146 168 LRR 5.

```

FT REPEAT 194 216 LRR 7.
FT REPEAT 217 240 LRR 8.
FT REPEAT 242 264 LRR 9.
FT REPEAT 266 288 LRR 10.
FT REPEAT 289 312 LRR 11.
FT REPEAT 313 336 LRR 12.
FT REPEAT 337 360 LRR 13.
FT REPEAT 361 384 LRR 14.
FT REPEAT 386 408 LRR 15.
FT REPEAT 409 432 LRR 16.
FT REPEAT 433 456 LRR 17.
FT REPEAT 458 478 LRR 18.
FT REPEAT 479 504 LRR 19.
FT REPEAT 506 529 LRR 20.
FT REPEAT 543 566 LRR 21.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66811 MW; DCD7637D94A5037C CRC64;

```

Query Match 9.2%; Score 344; DB 1; Length 603;
 Best Local Similarity 23.3%; Pred. No. 2.6e-17;
 Matches 141; Conservative 63; Mismatches 194; Indels 208; Gaps 16;

```

QY 8 LLLAWA-----GATATPVVPMVHVPCCACQIRWYPRSSYREATTVDN 56
DB 13 LIAFWALGCHLQGTDPGASDAE---GPCCPVACTCS-----HDDYTDLSVFC 61
QY 57 DFLTAVPAPLPAQGTOTLLQNSIVRV-----DQSLG 91
DB 62 SKNLTHLPDPIPVSTRALMDGNMSSIPSAQONSLDFLWQSGWLSRLEPQALG- 120
QY 92 LANLTLDLSQNSFSPARDCDFHALPOLLSLHEENQTRLE----- 133
DB 121 LQNTLYLHERRNLRLAVGLFTHPTSLASLSSNLLGLEGLPOGSLHMDLNQWN 180
QY 134 -----DHSEFAGIASQELIYLNNOIYRPAFRFSGISNLRLHLSNLRDAIDSRFEM 187
DB 181 SLVLPDPTVFOGSLNHELYLAGNKLTYLPALFCGLGELRELDLSRNLRSYKAVFVH 240
QY 188 LPNLELIMIGKNVDAT-----LDKN----- 208
DB 241 LPRQLKLYIDRLNLTAVAGAPLGMALKMLDLSHNRVAGLMEDTPGLGLHVLRLAHN 300
QY 209 -----FRPLANLRSVLVLAGMNLREISDYALBGLISLESLSFYDNOIARV----- 252
DB 301 AIASLRPRFTKDLHLEELQGNHRIQGERFEGLQLEVLITLNDQITTEVRVAFSG 360
QY 253 -----PRALQVYVGLKFLDLINKNPLORVGGDFANMLHKEGLGNM 295
DB 361 LFNVAVMNISGNCLSLPRERVQGLDKLHSLHESCLGHVRLHPTFAGLSGLRFLRD- 419
QY 296 BELVASIDKFLVNLPELTATLDTNNPRLSEIHRPRAFHLPOMETLMLNNALSAL----- 350
DB 420 NSISSIEBOSLAGLSLELDLITN-RLTHLPRLQFLQGLHLEYLLSYNOITLTSAEVL 478
QY 351 -----HQ-----QVESLPNQEVGLHGNPJR 372
DB 479 GLPQRAFWLIDISHNLETLAEGLFSSIGRVRLSLRNNLSQFFSPQGLERLMDLNPFMD 538
QY 373 CDCVIR-----WANNATGTRVFIEP-----OSTLCAEPPDLQRLPVREVPJR 414
DB 539 CSCPLKALRDFALQNPVGVPRVQTVCEGDDCOPVYTYNNITCAGPANSGLDLRDVSET 598
QY 415 EMTDHC 420
DB 599 HFV-HC 603

```

```

RESULT 5
ALS_MOUSE
ID ALS_MOUSE STANDARD; PRT; 603 AA.
AC P70389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS OR ALBS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID-10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129/SV;
RX MEDLINE-96413591; PubMed-8816745;
RA Boscclair Y.R., Seto D., Hasleh S., Hurst K.R., Ooi G.T.;
RT "Organization and chromosomal localization of the gene encoding the
RT mouse acid labile subunit of the insulin-like growth factor binding
RT complex."
RL Proc. Natl. Acad. Sci. U.S.A. 93:10028-10033(1996).
CC -1- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF
CC CIRCULATING IGFs TO THE TISSUES.
CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
CC IGF-1 OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; U66900; AAB17270.1; -.
DR MGD; MGI:107973; Igfals.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam; PF00560; LRR; 19.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 9.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 603
FT REPEAT 52 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 169 192
FT REPEAT 194 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 265 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 338 360
FT REPEAT 361 384
FT REPEAT 386 408
FT LRR 1.
FT LRR 2.
FT LRR 3.
FT LRR 4.
FT LRR 5.
FT LRR 6.
FT LRR 7.
FT LRR 8.
FT LRR 9.
FT LRR 10.
FT LRR 11.
FT LRR 12.
FT LRR 13.
FT LRR 14.
FT LRR 15.

```

```

SQ      SEQUENCE      567 AA;  63467 MM;  C48643AA73967A7D CRC64;

```

Query Match 9.1%; Score 340.5; DB 1; Length 567;
Best Local Similarity 24.6%; Pred. No. 4.2e-17;
Matches 132; Conservative 70; Mismatches 217; Indels 117; Gaps 15;

```

0Y      17 TATPVV---PMHPCPPQACQCR-----PWYT 42
      8 SAVPLLEAOPF---PCPKTCGVVRAAAGSGSVAHIAELGPTNLTHILFRMOGIL 65
0Y      43 PRSSYREATVVD---CNDLFTAVPA-----LPAG-----TQ 72
      66 RNHFSGGTVVQROMLSDSHSADPCTFNDVLVKTLRLTRNKISRLPAIIDKWLLE 125
0Y      73 TLLQSSIVYRDOSELGYLANLTELIDLSQNSFSDARCDPFAHPOLLSHLENOITRL 132
      126 QLFIDHNAIRLDQNLFOQLRNLQDELGNQNLSTFLPANFSSIREKLKLDLSRNNITLH 185
0Y      133 EDHSEFAGIASQELLYLHNQLYRIAPRAFSGLSNLRHLHNSMLRAIDSREEMLPNE 192
      186 PKGLIGAQVKLEKLILLYSNLTSDVSGILSNLGAITELRLRHNLRSVARGADRLGNLS 245
0Y      193 ILMIGAKKVVADIIDMNRPLANLSYVLAGNLRKISDYALBEGIQSIESSEFYDNOLARY 252
      246 SLTISGMLSESLPPALPLHWSVSRLTFENPLEELPDVLFGEAGAGREIMLNGTHILSTL 305
0Y      253 PRRALEQVPGKLFKDKNP-LQGVGGDANMMLHKEGLNNMEELVSDIKALVNLPE 311
      306 PAAAFRNLSIGLQGLTRNPRLSALPVGVOGLEKELVGLDHLN-NAIAELRDDALGLGH 364
0Y      312 LTKLDITNNPRLSFIHRAAFHNLPQMETLMLNNNAALSALAQGVESLSPNLQVGLGNP1 371
      365 LROYSLRHN-RLRALPRTLFRNLSSLESUGLEHNQLETLGVDVFAALPOLTYVLLGHNP4 423
0Y      372 RCDC---VIRM-----ANATGTVRFIEPOST-----CAEPDLOQLPY 408
      424 LCGDGLRFLQWLNRHNDILGRDEPPCGRPEPRASLSFELLQGPWCPDP----- 475
      409 REVEFRKMTQICPLISPRSPSLO-----YASGESVYLRCRALAEPEPIY 457
      476 RSLPDLPTERALLAEPVPSMLPNSWQCYTAOLVAREG-----PNNRLY 521

```

```

RESULT 7
SLIT_DROME
ID SLIT_DROME STANDARD; PRT; 1480 AA.
AC P24014;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE SLIT protein precursor.
GN SLI.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachytera;
OC Phagytoidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91099665; PubMed=2176636;
RA Rotherberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
RT SLIT: an extracellular protein necessary for development of midline
RT glia and commissural axon pathways contains both EGF and LRR
RT domains."?
RL Genes Dev. 4:2169-2187(1990).
CC -1- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND
CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR
CC MATRIX MOLECULES.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.

```

CC	-1	SIMILARITY:	CONTAINS 7 EGF-LIKE DOMAINS.
CC	-1	SIMILARITY:	CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
CC	-1	SIMILARITY:	CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC	-1	SIMILARITY:	CONTAINS 1 LAMININ G-LIKE DOMAIN.
CC	-----		
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	CC	use by non-profit institutions as long as its content is in no way	
CC	CC	modified and that statement is not removed. Usage by and for commercial	
CC	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	CC	or send an email to license@sib-ch)	
CC	-----		
DR	EMBL;	X53959; CAAB37910.1; ..	
DR	PIR;	A36665; A36665.	
DR	HSSP;	P00743; ICGF.	
DR	FlyBase;	FBgn0003425; sl1.	
DR	InterPro;	IPR000152; Asx_hydroxyl.	
DR	InterPro;	IPR000359; Cys_knot.	
DR	InterPro;	IPR000561; Egf_knot.	
DR	InterPro;	IPR000742; EGF_1like.	
DR	InterPro;	IPR001881; EGF_CA_2.	
DR	InterPro;	IPR001611; LRR.	
DR	InterPro;	IPR000483; LRR_Cterm.	
DR	InterPro;	IPR000372; LRR_Nterm.	
DR	InterPro;	IPR003592; LRR_out.	
DR	InterPro;	IPR003591; LRR_Typ.	
DR	InterPro;	IPR001791; Laminln_G.	
DR	Pfam;	PF00007; Cys_knot; 1.	
DR	Pfam;	PF00008; EGF; 7.	
DR	Pfam;	PF00054; laminln_G; 1.	
DR	Pfam;	PF00560; LRR; 17.	
DR	Pfam;	PF01463; LRRT; 4.	
DR	Pfam;	PF01462; LRNT; 4.	
DR	SMART;	SMO0041; CT; 1.	
DR	SMART;	SMO0179; EGF_CA_2.	
DR	SMART;	SMO0001; EGF_Like; 5.	
DR	SMART;	SMO0370; LRR; 4.	
DR	SMART;	SMO0082; LRRT; 4.	
DR	SMART;	SMO0013; LRNT; 4.	
DR	SMART;	SMO0369; LRR_Typ; 9.	
DR	SMART;	SMO0282; Lamg; 1.	
DR	PROSITE;	PS00010; ASX_HYDROXYL; 3.	
DR	PROSITE;	PS01185; CTCK_1; 1.	
DR	PROSITE;	PS01225; CTCK_2; 1.	
DR	PROSITE;	PS00022; EGF_1; 7.	
DR	PROSITE;	PS01186; EGF_2; 5.	
DR	PROSITE;	PS01187; EGF_CA_2.	
DR	PROSITE;	PS50025; LAM_G_DOMAIN; 1.	
KW	Neurogenesis;	Glycoprotein; Signal; Alternative splicing;	
KW	EGF-like domain;	Repeat; Leucine-rich repeat.	
FT	FT	SIGNAL	1 36
FT	CHAIN		37 1480 SLIT PROTEIN.
FT	REPEAT		99 122 LRR 1.
FT	REPEAT		123 146 LRR 2.
FT	REPEAT		148 170 LRR 3.
FT	REPEAT		171 194 LRR 4.
FT	REPEAT		195 218 LRR 5.
FT	REPEAT		220 246 LRR 6.
FT	REPEAT		321 344 LRR 7.
FT	REPEAT		345 368 LRR 8.
FT	REPEAT		369 392 LRR 9.
FT	REPEAT		394 416 LRR 10.
FT	REPEAT		417 440 LRR 11.
FT	REPEAT		522 545 LRR 12.
FT	REPEAT		546 569 LRR 13.
FT	REPEAT		570 593 LRR 14.
FT	REPEAT		595 617 LRR 15.
FT	REPEAT		618 641 LRR 16.
FT	REPEAT		643 666 LRR 17.
FT	REPEAT		678 701 LRR 18.
FT	REPEAT		720 743 LRR 19.
FT	REPEAT		745 764 LRR 20.

CC	-1-	TISSUE SPECIFICITY: PLATELETS AND MEGAKARYOCYTES.
CC	-1- PM:	THE N-TERMINUS IS BLOCKED.
CC	-1-	SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		the European Bioinformatics institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC		-----
DR	EMBL:	L11238; AAA03069.1; -;
DR	EMBL:	Z23091; CAA80637.1; -;
DR	HSSP:	P16473; IXOM.
DR	MIM:	173511; -;
DR	InterPro:	IPR001611; LRR.
DR	InterPro:	IPR000483; LRR_Cterm.
DR	InterPro:	IPR000372; LRR_Nterm.
DR	InterPro:	IPR003592; LRR_Out.
DR	InterPro:	IPR003591; LRR_Typ.
DR	Pfam:	PF00560; LRR_14.
DR	Pfam:	PF01463; LRCP; 1.
DR	PRINTS:	PRO0019; LEURICHRPT.
DR	SMART:	SMO0370; LRR; 1.
DR	SMART:	SMO0082; LRCT; 1.
DR	SMART:	SMO0013; LRNT; 1.
DR	SMART:	SMO0369; LRR_Typ; 9.
KW	Platelet;	Transmembrane; Glycoprotein; Blood coagulation;
KW	Signal;	Leucine-rich repeat; Cell adhesion; Signal.
FT	SIGNAL	1 16 POTENTIAL.
FT	CHAIN	17 560 PLATELET GLYCOPROTEIN V,
FT	DOMAIN	17 523 EXTRACELLULAR (POTENTIAL).
FT	TRANSERM	524 544 POTENTIAL.
FT	DOMAIN	545 560 CYTOPLASMIC (POTENTIAL).
FT	REPEAT	73 96 LRR 1.
FT	REPEAT	97 120 LRR 2.
FT	REPEAT	122 144 LRR 3.
FT	REPEAT	145 168 LRR 4.
FT	REPEAT	169 192 LRR 5.
FT	REPEAT	194 216 LRR 6.
FT	REPEAT	217 240 LRR 7.
FT	REPEAT	241 264 LRR 8.
FT	REPEAT	266 288 LRR 9.
FT	REPEAT	289 312 LRR 10.
FT	REPEAT	314 337 LRR 11.
FT	REPEAT	338 361 LRR 12.
FT	REPEAT	362 385 LRR 13.
FT	REPEAT	386 409 LRR 14.
FT	REPEAT	51 51 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	181 181 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	298 298 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	312 312 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	385 385 N-LINKED (GLCNAC. . .)
FT	CARBOHYD	499 499 N-LINKED (GLCNAC. . .)
FT	CONFLICT	73 74 MT -> TK (IN REF. 2).
FT	CONFLICT	109 109 K -> T (IN REF. 2).
FT	CONFLICT	130 130 D -> W (IN REF. 3).
FT	CONFLICT	136 138 GID -> PGG (IN REF. 3).
FT	CONFLICT	209 209 L -> I (IN REF. 2).
FT	CONFLICT	267 267 N -> H (IN REF. 3).
FT	CONFLICT	327 327 L -> I (IN REF. 2).
FT	CONFLICT	478 478 P -> G (IN REF. 2).
FT	CONFLICT	509 509 P -> D (IN REF. 2).
Q	SEQUENCE	560 AA; 60959 MW; B1CDB04AFAF7115 CRC64;

QY	9	LLAMVAGTATVPVPMVHVPPOCAQOIRWYTPRSSRYEATRVDCNDLF-LTAAVPAL	67
QY	9	LLAMVAGTATVPVPMVHVPPOCAQOIRWYTPRSSRYEATRVDCNDLF-LTAAVPAL	67
Db	6	LLCAVLGILLRAP-----FPCPPACK-----VFNDAAQCCSGDVARISAL--GL	48
QY	68	PAG-TOTLLLOSNIVRYDQSEGLAYLANLTIELDSQNSFSPARCDCHAPOLLSLHLE	126
Db	49	P7NLTHILLFCMGKGVQSQSFSG-MYVLQRLMISDSHISAVNGTFSDLTKLTLSLR	107
QY	127	NOLTRLEDHSPAGIASLOELYLNHNOLYRIAPRAFGSLNLLRIHLNLSNLRAIDSFWE	186
Db	108	NKITHLPALLDKVWLLLEOLFEDHNAALRGIDQNNFQKLVNLQELIALNODLDFASLFT	167
QY	187	MLPNEILMIGCNVVDALIDMNFPRPLNLSVLAAGNNLEISDYALEGLQSELSFYD	246
Db	168	NLENKTLIDSGNNLITPHPKGLGAQAOKLERLLHSNRLVSLDSGLNLSIGALTTELOFHR	227
QY	247	NOLARVPRALEQVPGKLFEDLNK-----NPLQRYGPDFA	282
Db	228	NHISIAAGADRLPRNLSLSLTSXNHILAFPSALFELSHNLTILITFENPLAELPGVLFG	287
QY	283	NMLHLKEGLNMMEEVYSIDKFAVLNLPETLKTIDTNNPRLSFTHPRAFHHPOMETIHL	342
Db	288	EMGQELMELNLR-TQLRTLPAAFRNLSRLRYLGVLTSPRLSALPGCAFQGLGELYAL	346
QY	343	NNNLSAL-----HQ-----T	354
Db	347	HSNGTLTAPDGLGLGCKLRQVSLRRNRRLRALPRALFRNLSLESYOLDNNOLETLPGDV	406
QY	355	VESLPNLOEVLGHNPJRCDC---VTRMANATGVTRFTEPGSTGLAEPDQLRLPVRE	410
Db	407	FGALPRLETVLLGHNNSRCDGGLGFLGWLROHLGLVGGEPPR--CAGGAHAGLPYMA	464
QY	411	VPFREMTHCLPLISPSFPSLQVAGSESWYLHCRAIAPEPEITY-WTP	460
Db	465	LPGDA--EC---GPRGPPRPAADSSSEAPVH-PALAPNSSEPPWMAQ	509
RESULT	9		
TLR3_MOUSE	TLR3_MOUSE	STANDARD;	PRT; 905 AA.
AC	099MB1: 0912M4;		
DT	01-MAR-2002 (Rel. 41, Created)		
DT	01-MAR-2002 (Rel. 41, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Toll-like receptor 3 precursor.		
GN	TLR3.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=129/SV;		
RX	PubMed=11607032;		
RA	Alexopoulos L., Holt A.C., Medzhitov R., Flavell R.A.;		
RT	"Recognition of double-stranded RNA and activation of NF-kappaB by		
RT	Toll-like receptor 3."		
RL	Nature 413:732-738(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BALB/c x NIH; TISSUE=Macrophage;		
RA	Appelquist S.E., Ljunggren H.G.;		
RT	"Molecular cloning of mouse Toll-like receptor 3 cDNA."		
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.		
CC	-!- FUNCTION: Participates in the innate immune response to microbial		
CC	agents. May be involved in the recognition of ds-RNA. Acts via		
CC	MyD88 and TrAfe, leading to NF-kappa-B activation, cytokine		
CC	secretion and the inflammatory response (By similarity).		
CC	-!- SUBUNIT: Binds MyD88 via their respective TIR domains (By		
CC	similarity).		
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).		
CC	-!- TISSUE SPECIFICITY: Highly expressed in lung. After		
CC	intraperitoneal injection of lipopolysaccharide, highly expressed		

FT	DOMAIN	359	379	SER/THR-RICH.
FT	CARBOHYD	53	53	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	90	90	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	98	98	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	207	207	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	245	245	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	327	327	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	495	495	N-LINKED (GLCNAC. . .) (POTENTIAL) .
SQ	SEQUENCE	536 AA;	58649 MW;	D524011243C9B2A6 CRC64;
Query Match		8.0%;	Score 288;	DB 1; Length 536;
Best Local Similarity		25.2%;	Pred. No. 4,6e-14;	
Matches 106;	Conservative	64;	Mismatches 192;	Indels 58; Gaps
QY	30	PROGACQIRPWTTPRSTSYREATVYDNDLFTRAVPALPAGITQLLGNSNIVYRDQSEL	89	
DB	2	PMGCDCEFTV-----EVFCSDBELATVPDIPPTKNIIVFTSFTLETRAF	48	
QY	90	GYLANTELDELSONSFSDARDCDFNALPOLLISLHEENQUTRLDEHSFAGLSLOETLYN	149	
DB	49	GSNPNLTKVYFLDQLOCGRPDPAFGSLPRLEDEVTGSSFLNLTINIFENLTLGKLTITN	108	
QY	150	HNOLYRIAPRASGSLNLIHLNLSNLRAIDSWEFEMLPNLEITMIGNKYDALIDNMF	209	
DB	109	FNNLEALPEGLRQHLAALESLHLOGNOQLPRRLPPLTHLKTMLNAQNLNLAQPEELF	168	
QY	210	REPLANRSYLVLGMMNRELSDVALDELQLESLSPYDNLAVPRPALQVYGLKFLDIN	269	
DB	169	HPLTSLQTLTKLSNNALSGLPQGVFGKLSQLEFLDSNNISLPPQVFSQLCERLWQ	228	
QY	270	KNPLOVGPDEPANMLHLKEIGLN-NMELVSIIDKVALNLEPRLKDTTNNPRLSFIHP	328	
DB	229	RNATIHPLPSIFASIGNLFLSLQNNMLRVLPAGLFA--HTFCYVGLSLITHN-QLETYTE	285	
QY	329	RAFPHLPMQETMLMNNNA-----LSALHOQTVBSLPNLOEV	364	
DB	286	GTFPHLSNLSRLSMLSYNMAITHLPAGIFRDLLETAKLYGSSNNLFTALHPALFQNLKLELL	345	
QY	365	GLHGPIR-----CDCYIRANMTGRVRRTEPQSTLCAPPPLOLRP	407	
DB	346	SLSKNQLTTLPASSTPTTTCSTWPCVTWPGSATPAGLPLOLAAVHRSAPRHPDLRCP	405	
RESULT 13				
TLR3_HUMAN				
ID	TLR3_HUMAN	STANDARD:	PRT:	904 AA.
AC	O15455;			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Toll-like receptor 3 precursor.			
GN	TLR3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96118556; PubMed=9435236;			
RA	Rock F.L., Hardiman G., Timans J.C., Kastelein R.A., Bazan J.F.;			
RT	"A family of human receptors structurally related to Drosophila			
RT	Toll."			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).			
CC	-I- FUNCTION: Participates in the innate immune response to microbial			
CC	agents. May be involved in the recognition of ds-RNA. Acts via			
CC	MYD88 and TRAF6, leading to NF-kappa-B activation, cytokine			
CC	secretion and the inflammatory response (By similarity).			
CC	-I- SUBUNIT: Binds Myd88 via their respective TIR domains (By			
CC	similarity).			
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).			
CC	-I- TISSUE SPECIFICITY: Expressed at high level in placenta and			

[illegible]

DB 330 IDPSLEHLTSLCFLWLSNRCLTFEARRLGSLPCLMLDLSDNALLETLELGARALGSLR 389
QY 276 -----VGPGEFANMLHKL-----GNNMEELY 299
DB 390 TLLGNALRDLPYTFYFANLASIQRLNLOGNRVSPGGPDEPSPGCVAFSGITSLASLS 449
QY 300 SIDK-----FALVNLPELTJKIDITNNP-----321
DB 450 LYDNEIELLRAGAFHTP-LTELDLSSNPGLEVAATGALGSLASLEVALOGNLMVLOY 508
QY 322 -----RLSIHPRFHHLP-----QMETLMLNNALSLHOOQVBSL-PNIOEVG 365
DB 509 DLPCFICLRNLNAENR-ISHLPAMTQAVSLEVLDLRNNSFILLPGSAMGLETSIRRLY 567
QY 366 LHGNPIRCDCVIRW--ANATGRFRFIEPOSTJC 397
DB 568 LQGNPLSC-CGNGWLAQAHQGRVDAVDATODLIC 600
RESULT 15
CHAO_TRICA STANDARD; PRT: 782 AA.
AC P82963; 16-0CT-2001 (Rel. 40, Created)
DT 16-0CT-2001 (Rel. 40, Last sequence update)
DT 16-0CT-2001 (Rel. 40, Last annotation update)
DE Chaoptin (Photoreceptor cell-specific membrane protein) (Fragment).
GN CHP
OS Tribolium castaneum (Red flour beetle).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiforma; Tenebrionidae; Tribolium.
OX NCBI_TaxID=7070;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GA-1;
RA Nie W., Stronach B., Panganiban G., Brown S., Denell R.;
RT "Molecular characterization of Tc1alpha and the 3' end of the
RT tribolium homeotic complex.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR PHOTORECEPTOR CELL MORPHOGENESIS. MEDIATES
CC HOMOPHILIC CELLULAR ADHESION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR SURFACE OF R-CELL PLASMA
CC MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAOPTIN FAMILY.
CC -1- SIMILARITY: CONTAINS AT LEAST 21 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF32227; AA01654.1;
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_LTP.
DR Pfam; PF00560; LRR; 19.
DR PRINTS; PR00019; LEUCRICHPT.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00369; LRR_LTP; 14.
KM Glycoprotein; Membrane; Repeat; Leucine-rich repeat; Vision.
FT NON_TR 1
FT REPEAT 43 66 LRR 1.
FT REPEAT 67 90 LRR 2.
FT REPEAT 93 116 LRR 3.
FT REPEAT 117 140 LRR 4.
FT REPEAT 141 164 LRR 5.
FT REPEAT 165 188 LRR 6.
FT REPEAT 224 247 LRR 7.

FT REPEAT 249 272 LRR 8.
FT REPEAT 273 296 LRR 9.
FT REPEAT 297 320 LRR 10.
FT REPEAT 321 343 LRR 11.
FT REPEAT 344 367 LRR 12.
FT REPEAT 370 391 LRR 13.
FT REPEAT 395 418 LRR 14.
FT REPEAT 419 444 LRR 15.
FT REPEAT 446 469 LRR 16.
FT REPEAT 491 513 LRR 17.
FT REPEAT 514 537 LRR 18.
FT REPEAT 539 562 LRR 19.
FT REPEAT 615 638 LRR 20.
FT REPEAT 640 663 LRR 21.
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 782 AA; 88561 MW; 311460B2D4527917 CRC64;
Query Match 7.5%; Score 279; DB 1; Length 782;
Best Local Similarity 24.8%; Pred. No. 1,9e-12;
Matches 145; Conservative 94; Mismatches 233; Indels 112; Gaps 23;
QY 47 YRRATTVDCNDLELT-AVPPALPAGTQTLLQSNSTVVRQDSGLGLANTLELDISONSF 105
DB 5 HRRQRFVDIGDSLTLTKLTHAL-----SSVQNFPSDAIKILNLEEDLSNNRL 54
QY 106 SPARDCEFHLPOLLSTLHLEENOLTRLEDHSFAG--IASLOELYLNHNOLYRIAPAFSG 163
DB 55 RNVDPNSFFHPLSLKKVNHQDNTIEIHNGTQGDHRLDTEVFEFSNVRYNQHTFAD 114
QY 164 LSNLRHLNLSNLRRAIDSRWEMLPNLEILMIGKRVDAIIDKRNPRPLANTSLVLAGM 223
DB 115 LLOLEIHLDNKRIESLERAFANLKLKRLMLKGNKIKTIAVEPTQNPPELDLDAVY 174
QY 224 NIREISDVALEGQSLSELSFYDN-----QLARPRALREVPCL-----KFLDLNK 270
DB 175 STISLDFNITFDVGSIG--MFHVNSSHKLINLVAVSVPFQCTGLGIGQNIKVLDLSE 232
QY 271 NPLQRYPGDFAFM-LHLKELGINMEELVSIDKFALVNLPELTJKIDITNNP--RLSFTI 327
DB 233 NNITSVAKQFFRPVELSLMQLYLGH-NKLLNMTKDLFGMMPHLQVLDLSHNSIYELDF-- 289
QY 328 PRAFHHLPMQETLMLNNALSLHOOQVBSLPNLOEVLGHNPTRC--DCVIRMANATGT 385
DB 290 -DTFRNTRKQLQWLDIYSHNRISLPINDLFFGLNLRIVDESHNRKLSLPNLR--ETG- 344
QY 386 RVRFIEPOSTLCAEPDRLRLPVREVPFREMDDHCLPLSPSPFSFSLQVAGSGSVNLIC 445
DB 345 -----LRLDVSSHNLGK-----LPL-----TSLSLASQTL----- 371
QY 446 RALAPEPEIYVTPAGLRLTPAHAGR--RYRVPEEGTELELRVTAEEGLTYCV---AQ 500
DB 372 -----SELDLSNKSISL-----SHGQLARFKCLSMULSYRLOQIDAGTKGIPRLAS 422
QY 501 NLVGADITVSVVGRALLQPGRDEGQLELRVQETHPYHILLSWV---TPPNTVSTNLT 557
DB 423 INLGHNSQTLLEINLSF-----QGLEYTLILANDVNSLSQVPAISPALSTPNLSLSLA 474
QY 558 MSS-----ASSLR--CGGATALARLPRGHSYNITPRL 588
DB 475 FNSLPTVALEVAGNISLRLYINLDYNDLSAVPIVYHSLTELRHL 518

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:36:13 / Search time 34.24 Seconds

(without alignments)
3602.377 Million cell updates/sec

Title: US-09-905-088a-245

Perfect score: 3732

Sequence: 1 MRLVAPLLAWAGATATV.....RKLPSSSEGTLLPPLSQNS 713

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mmc:*
8: sp_mammal:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaeoplasmid:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1695	45.4	730	4	09P231
2	1687	45.2	716	11	061809
3	1656.5	44.4	705	4	043377
4	1656.5	44.4	708	4	09H3W5
5	1649.5	44.2	707	11	09ESY6
6	1648.5	43.6	718	13	P97860
7	1625.5	43.6	718	13	073675
8	875.5	23.5	431	4	09H8V1
9	755.5	20.2	273	4	09H8V1
10	494.5	13.3	614	11	09D1T0
11	492.5	13.2	614	4	096FE5
12	490	13.1	614	6	09N008
13	468.5	12.6	606	4	09B220
14	442	11.8	1091	11	P70193
15	438.5	11.7	640	4	09HC23
16	432.5	11.6	540	5	09V053

17	431	11.5	1094	4	09HYB8	09HYB8 homo sapien
18	430	11.5	1093	4	096JAI	096JAI homo sapien
19	427.5	11.5	532	5	096671	096671 drosophila
20	420.5	11.3	628	4	09BPTN	09BPTN homo sapien
21	407	10.9	733	5	024250	024250 drosophila
22	407	10.9	737	5	09VU51	09VU51 drosophila
23	406.5	10.9	719	4	096NT6	096NT6 homo sapien
24	401.5	10.8	737	5	0965M3	0965M3 caenorhabdi
25	401.5	10.8	881	5	0965M2	0965M2 caenorhabdi
26	401.5	10.8	1447	5	016779	016779 caenorhabdi
27	399.5	10.7	811	4	075139	075139 homo sapien
28	399	10.7	78	11	061974	061974 mus musculu
29	398.5	10.7	789	6	09BE71	09BE71 macaca fasc
30	395.5	10.6	673	11	09CZT5	09CZT5 mus musculu
31	392.5	10.5	792	4	09UDT7	09UDT7 homo sapien
32	388.5	10.4	832	4	09ULH4	09ULH4 homo sapien
33	386	10.3	492	11	099KT6	099KT6 mus musculu
34	386	10.3	653	4	09HBW1	09HBW1 mus musculu
35	385	10.3	788	11	09CYK3	09CYK3 mus musculu
36	378	10.1	1065	4	096A85	096A85 mus musculu
37	377	10.1	1065	4	094898	094898 mus musculu
38	374.5	10.0	700	4	09P244	09P244 mus musculu
39	368	9.9	907	11	0921P4	0921P4 mus musculu
40	361.5	9.7	542	5	09N4G6	09N4G6 caenorhabdi
41	361.5	9.7	809	11	09DBY4	09DBY4 mus musculu
42	361	9.7	1173	5	09V7J8	09V7J8 drosophila
43	360.5	9.7	1531	11	088279	088279 mus musculu
44	360	9.6	473	11	099P18	099P18 mus musculu
45	360	9.6	907	4	075473	075473 homo sapien

ALIGNMENTS

RESULT 1
ID 09P231 PRELIMINARY: PRT: 730 AA.
AC 09P231
DI 01-OCT-2000 (TREMBLrel. 15, Created)
DI 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DI 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DI KIAA1497 PROTEIN (FRAGMENT).
DI KIAA1497.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20277482; PubMed-10819331;
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Chara O.,
RT "Prediction of the coding sequences of unidentified human
RT genes. XVII. The complete sequences of 100 new cDNA clones from brain
RT which code for large proteins in vitro."
RL DNA Res. 7:143-150(2000).
DR EMBL, AB040930; BA96021.1;
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR.
DR InterPro: IPR000372; LRR_Cterm.
DR InterPro: IPR003592; LRR_Out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF00560; LRR_9.
DR Pfam: PF01463; LRRCT_1.
DR Pfam: PF01462; LRRNT_1.
DR SMART: SK00060; FN3_1.
DR SMART: SK00408; IGC2_1.
DR SMART: SK00370; LRR_4.
DR SMART: SK00082; LRRCT_1.
DR SMART: SK00013; LRRNT_1.

QY 619 ----TSCHRALGDRPGILALALAVLL 641
 DB 626 HETSTALAAVWGSMFAVISLAIYI 652

RESULT 3
 ID 043377 PRELIMINARY; PRT; 705 AA.
 AC 043377;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PROBABLE LEUCINE-RICH REPEAT PROTEIN.
 GN RGL18D07.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FIBROBLAST;
 RA Murray J., Langston Y., Clarke C.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC004142; AAC02752.1;
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003598; Iq_C2.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00041; In3; 1.
 DR Pfam: PF00047; In3; 1.
 DR Pfam: PF00560; LRR; 9.
 DR Pfam: PF01463; LRRCT; 1.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PRO0019; LEURICHRPT.
 DR SMART: SM00408; IGC2; 1.
 DR SMART: SM00370; LRR; 4.
 DR SMART: SM00082; LRRCT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_typ; 1.
 DR Immunoglobulin domain.
 SK SEQUENCE 705 AA; 79049 MW; 984066693FABCI4 CRC64;

Query Match 44.4%; Score 1656.5; DB 4; Length 705;
 Best Local Similarity 47.2%; Pred. No. 6e-118;
 Matches 333; Conservative 109; Mismatches 217; Indels 47; Gaps 5;

QY 17 TATVTVVFMHVCPCQCAQIRPWTPRSSYREATTVCNDLFLRVAPALPAGTQITLL 76
 DB 14 TTVVAVVKKKDCPCPLCTCCTCIRPWTFRSTYMEASTVDCNDGLTFPAPLPAANOIILL 73

QY 77 QSNSTIVRVDSGLYANITELDLSONSFSDARDCDFHLPOLLSLHLENOQLTLEDHS 136
 DB 74 QTNNTAKIEYS TDEPPVNLTGDLDSQNNLSVTNINVKKMPOLLSVYLEENKTELPEKC 132

QY 137 FAGLASLOELYNQLYRIAPRAFSGLSNLRHLNSNLRRAIDS RFEMLPNIETIMI 196
 DB 133 LSELSNLOELYINHMLSTISPGAFIGLHNLRLHLSNRLQNSKMPDAPNIEIIMI 192

QY 197 GGNKVDATLDNFRPLARLSYLAGMNLRESDVALSGLOSLESLFYDNOALAVPRRA 256
 DB 193 GGNPTIRIKDNKFKPLNLRSLVYAGINLTETPDNALVGLMLESISFDNMLIVPHYA 252

QY 257 LEQVPGKFLDLNKNPLQVGVGFANMLHLKELGNNNEELVSTDKFALVNLPELTAK 316
 DB 253 LQKYVNLKFLDLNKNPIKIRIRGSDSNMLHLKELGINNPELISIDSLAVDNLPLRLIE 312

QY 317 ITNPRLSFTHPRAFHLLPQMETLNNNALSAHQQVESLPNLOEVLGHNPIRCDCV 376
 I|||||:||||| I|:|||||:||||| I|:|||||:|||||

DB 313 ATNPRLSYIHPNAFRLPKLESIMNSMLSAHYHGIESLPNKEISHSNPIRCDCV 372
 QY 377 IRANANAGTRVRFTPEOSTLCAEPDLORLPVREVPREMTDCLPLISPRSPSIOVA 436
 DB 373 IRWMNMKNKTNIRFEMPEPSLFCVDPPEFGQVNRQVHFEDMMEICLPLIAPSPFSNLNVE 432

QY 437 SGESMVLHCRALAEPEPEIYVMPAGLRTPAHAGRRYRYPESTLELRVTAEGILYT 496
 DB 433 AGSVSFHCRALAEPEPEIYVMPAGLRTPAHAGRRYRYPESTLELRVTAEGILYT 492

QY 497 CVAQNLVADTKVSVVGVGALLPGRGDEGQLELRVQETHPYHILSWTPPTVSTNL 556
 DB 493 CIATNLVGLDKSMIXVDSFQ---DNNGSLNIKIRDIQANSVLSMASSKILKSSV 549

QY 557 TWSSASSLRQCATALALRPRGTHSYNITRLQATEYMACLOVAFADAHNOLACWARTK 616
 DB 550 KWTAFVKTENSHAAQSARIPSDVKVYMLTHNPSTEXKICIDIFTYIKRKKCVANTTK 609

QY 617 -----EATSCHRALGDRPGILALALAVLLAAGLANHLETGQPRKGVGRR- 663
 DB 610 GLHPDQKEYKKNNTTLMACLGILGIVICLISCL-----SPENNCDSGHS 657

QY 664 -----PLPPMAAFWGWGAPSVRVVSAPLVLPWN 691
 DB 658 YVRNYLQKPTFPALGELYPPLINEMEGKKEKSTSLKAKATYIGLPTN 703

RESULT 4
 ID 09H3W5 PRELIMINARY; PRT; 708 AA.
 AC 09H3W5;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HYPOTHETICAL 79.4 KDA PROTEIN (NEURONAL LEUCINE-RICH REPEAT PROTEIN-3).
 GN DKFPZ76LX2424 OR NLRR-3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=AMYGDAIA;
 RA Bloecher H., Boecher M., Brandt P., Nemes H.W., Weil B., Wiemann S.;
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB060967; BAB47184.1;
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003598; Iq_C2.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00041; In3; 1.
 DR Pfam: PF00047; In3; 1.
 DR Pfam: PF00560; LRR; 9.
 DR Pfam: PF01463; LRRCT; 1.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PRO0019; LEURICHRPT.
 DR SMART: SM00060; FN3; 1.
 DR SMART: SM00409; Iq; 1.
 DR SMART: SM00408; IGC2; 1.
 DR SMART: SM00370; LRR; 4.
 DR SMART: SM00082; LRRCT; 1.

```

Db 483 ICGITPREGGLTICATNVLGADLSIMKVGGEVPO---DNGSLNLIKIRIDIRANSVLY 539
QY 544 SVTPPNTVSTNLTWSSASSLRQGATATALRLPRGTHSYNITRLLOATEYMACLOYAFAD 603
Db 540 SWKANSKILKSSVKMTAFKTEDSQAAQSARIPSDVKVNLTLTKSTEEKICIDITPTIY 599
QY 604 AATQALACVART-----KEATSCH---RALGDRPGLIAIALAVLLAAGLAHLGT 652
Db 600 QKSRQCVNVTTKSLSHDEKENGKSHTEVACVCGILGIIYVCL-----FCVSGOEGN 653
QY 653 GQPRKGVGRRLPPAMAFMGNSAPSVRYSAFLVLPMPNGRLPSSSE 701
Db 654 CENHSYTVNHCHKPTLAF-----SELYPLINLMESSKKEKPA5LE 694

```

```

RESULT 6
P97860 PRELIMINARY; PRT; 707 AA.
AC P97860.
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LECICINE-RICH REPEAT PROTEIN PRECURSOR (FRAGMENT).
GN LRRN3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96233817; PubMed=9011764;
RA Taniguchi H., Tohyama M., Takagi T.;
RA "Cloning and expression of a novel gene for a protein with leucine-
RL Brain Res. Mol. Brain Res. 36:45-52(1996).
DR EML; D9802; BAA08622.1; -.
DR MGD; MGI:106036; Lrrn3.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003598; Iq_c2.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; fn3; 1.
DR Pfam; PF00560; LRR; 10.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00682; LRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 1.
KM Immunoglobulin domain; Signal.
FT SIGNAL 1 20
FT CHAIN 21 >707 POTENTIAL.
FT NON_TER 707 707
SQ SEQUENCE 707 AA; 79156 MW; FAETC1573DDD165B CRC64;

```

Query Match 44.28; Score 1648.5; DB 11; Length 707;
 Best Local Similarity 47.78; Pred. No. 2.5e-117;
 Matches 328; Conservative 111; Mismatches 225; Indels 23; Gaps 5;

```

QY 17 TATPVVPMHPCPPQCCQIRPMTPTSSYREATTVCNDLFLAVPAPLAGQTILL 76
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 TTVQALDKKDCFCOLCEIRPFPTSTIMEASTVDCNDGLNLPARLPADTOILL 76

```

```

QY 77 QSNISYRVDSSELYLANLTETDLSQNSFSDARDCHALPOLLSTLSENOITREDDHS 136
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 QTNINRIEHS-IDFPVNLGIDLSQNNLSVTNINQKNSQLSTYLENKLTLEPEC 135
QY 137 FAGLASLOEYLYNNOLYRIAPAFSGLSNLRHLHNSNLRRAIDSRWFEMLPLETIM 196
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 LVGLSNLOELYNNHNLSTSPGAFIGLHNLRLHNSNLRQMINSQWFDALPNLEITML 195
QY 197 GGNKVDAILDMNRPRLANRLSYLAGMNLREISDYALEGQSLESFYNQIARPPRA 256
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 GDNPIIRIKDMNFOPLVKLSLVIAIGINTEIPDDLAGLENLESISFYDNRLSKYPOVA 255
QY 257 LEQVPLKEFLDKNKPLOVRGPDFANMLHLKELGANNMEELSIDKFAVLNLPETIKD 316
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 LQKAVNLKFLDLNKNP INTRRGDFSNMLHLKELGINNPELVSIDSIAVDNLPDLKTE 315
QY 317 TNNPRLSFIHPRAFHLLPOMETIMLNNSALHQOYVESLPNLOEVLGAPRDCV 376
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 ATNNPRLSYIHPNAFRLPLESILMTNTNALSALYHGTEISLPNLEKISHSNPRDCV 375
QY 377 IRMANNTGTRVREIEPQSLCAEPDQLPVREVPFRENTDCLPLISRPSPSLQVA 436
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 IRWINNKNTRINREMPDLSFCYDPPEQONVQVFRDMMEICPLIAPESPSDLVE 435
QY 437 SGEVNLHCRALAEPEPEIYWTAGLRTPAHAGRRYRVPEGTLELRVTAEEAGLYT 496
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 ADSYVSLHCRATRAEPREIYMTIPSSKLLPNTMRKFYVHSGTIEITPKEGLYT 495
QY 497 CVAQNLVGADTVSVVGRALLQGRDEGQGLELRVOETHPYHILLSWTPPNTVSTNL 556
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 496 CIATNLVGADLSIMIKYGVSVPO---DNGSLNLIKIRIDIRANSVLYSVKASKILKSRV 552
QY 557 TWSSASSLRQGATATALRLPRGTHSYNITRLLOATEYMACLOYAFADANTQALCVARTK 616
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 553 KWTGEVKTEDSHAQASARIPSDVKVNLTLTKSTEEKICIDIPFYOKSRQCVNVTTK 612
QY 617 EATSCHRALGDR-----PGLIATIALAVLLAAGLAHLGTGQPRKGVGRRLPPA 668
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 613 SLEHCKEYGNHITVYVACVGLIGITGVYVCLF---SCVSQBSGSGEHSYAVNHCHKPA 669
QY 669 WAFMGNSAPSVRYSAFLVLPMPNGRK 695
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 670 LAF-----SELYPLINLMESSKE 688

```

```

RESULT 7
ID 073675 PRELIMINARY; PRT; 718 AA.
AC 073675;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NEURONAL LECICINE-RICH REPEAT PROTEIN.
GN XNLR-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99070063; PubMed=9852961;
RA Hayata T., Uochi T., Asashima M.;
RA "Molecular cloning of XNLR-1, a Xenopus homolog of mouse neuronal
RT leucine-rich repeat protein expressed in the developing xenopus
RL nervous system."
RL Gene 221:159-166(1998).
DR EML; AB014462; BAA28530.1; -.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003598; Iq_c2.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.

```

DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PLAC1006239 PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Ito T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugeno S., Ishibashi T., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirooka S., Ishii S., Kawai Y.,
 RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,
 RA Masuno Y., Kanehori K.;
 RT "NEO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK001991; BAA92025.1; -;
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR; 6.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PRO0019; LEURICHRPT.
 DR SMART; SM00370; LRR; 2.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 1.
 DR NON_TER 273
 FT SEQUENCE 273 AA; 30832 MW; B59F070A6D7D11FF CRC64;

Query Match 20.2%; Score 755.5; DB 4; Length 273;
 Best Local Similarity 58.8%; Pred. No. 1e-45; Mismatches 66; Indels 1; Gaps 1;
 Matches 151; Conservative 39; Mismatches 66; Indels 1; Gaps 1;
 QY 17 TATVPMVPHVPCPOCAQIRPWYTPRSSYREATTDCNDLFTAVPAPAGTOLL 76
 DB 17 TLVAVDKKVDKPCRLCTGCEIRWETPRSTIMEASVDSGLLTFPRLPANTOILL 76
 QY 77 GNSNIVRDSQELGYANTTELDLSNFSVDACDFHALPOLLSHLEENOTRLEDS 136
 DB 77 QNNIAKLEYS-TDEPVNLTGLDLSQNNLSVTYNVKKMPOLLVSYLEENKTEPEKC 135
 QY 137 FAGLASLQELYNHNOQYRIAPRAFGSLNLRHLNSNLRAIDSRWEMPNLEIIM 196
 DB 136 LSELNQLQELYNHNLSTISPGATIGLNLRLHLNSNRLQWINSKWDALPNLEIMI 195
 QY 197 GKNKYDALIDNFRPLANRSLVLAGMNLREISDYALEGLQSLSTSFYDNOCLARVPR 256
 DB 196 GNPPIRIKIDNMFPLINRSLVINGINLTLPDNLVLENLSTSFYDNLRIKYPHYA 255
 QY 257 LEQVGLKFLDKNKP 273
 DB 256 LQKVNLKFLDKNKP 272
 RESULT 10
 Q9D1T0 PRELIMINARY; PRT; 614 AA.
 AC Q9D1T0;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ADULT MALE TESTIS CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
 DE CLONE 4930471K13, FULL INSERT SEQUENCE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arawaka T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
 RA Kuehl P., Lewis S., Matsuo Y., Mikaldo I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK027262; BAB32403.1; -;
 DR HSSP; P23945; 1XUN.
 DR InterPro; IPR003599; I9.
 DR InterPro; IPR003598; I9_C2.
 DR InterPro; IPR003600; I9_Like.
 DR InterPro; IPR003006; I9_MHC.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR Pfam; PF00047; I9; 1.
 DR Pfam; PF00560; LRR; 9.
 DR Pfam; PF01463; LRRCT; 1.
 DR Pfam; PF01462; LRRNT; 1.
 DR SMART; SM00409; I9; 1.
 DR SMART; SM00408; I9C2; 1.
 DR SMART; SM00410; I9_Like; 1.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 9.
 KW Immunoglobulin domain.
 SQ SEQUENCE 614 AA; 69100 MW; 41CFP40C21335681 CRC64;

Query Match 13.3%; Score 494.5; DB 11; Length 614;
 Best Local Similarity 25.6%; Pred. No. 2.9e-29;
 Matches 173; Conservative 97; Mismatches 251; Indels 155; Gaps 18;
 QY 1 MRLVAPLLAM-----VAGTAVPVVPMVHPCPOCAQIRPWYTPRSSYREA 50
 DB 6 MRSMPSPLACQPIILLVIGSVLSGSAF-----GCPPRECS-----AGD 46
 QY 51 TVVDCNDLFTAVPAPAGTOLLQSNISYVRSQELGYANTTELDLSNFSVDAD 110
 DB 47 RAVLCHRRKRFVAVPBGIPETRLDIGNKRIITLNDDEFASPHLEELNENIYSAVP 106
 QY 111 CDFH-----ALPOLLSHLEENOTRLEDSFAGLASLOEL 146
 DB 107 GAFNNLNFNRTIGLSNRKLPLGVFTGLSNLTIDISENKIVILLIDYFODLYNLKSL 166
 QY 147 YLNHQLRIAPRAFGSLNLRHLNSNLRAIDSRWEMPNLEIMIGNKRYDAIID 206
 DB 167 EVGDNDIVYISHRASGLNLSQLTLEKCNLITSIFTEALSHLGLIVLRLHNTAIND 226
 QY 207 MNERPLANRSLVLAGM-NLREISDYALEGLQSLSTSFYDNOCLARVPRALQVGLKF 265
 DB 227 YSFKRLRYKLYEISHWPYLDTPNCLYGL-NLSTLSTHCNLTAVPYLAVRHLYVLYLF 285
 QY 266 IDLNKNPQVRGPDGFANMLHLKELGLNNMEETVSDIKFALVNLDELVTLDITNNPRISF 325
 DB 286 LNLSTNPIGTIE-----GSMTH-----ELLKQELQVY-----GGQLAV 319

```

QY 1 MLLVAPLLIAW-----VAGATATVPVYVWVPCPOCAQOIRPWYTPRSSYREA 50
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 VRMSPLSLACWQPIILLVGLVSGSAT-----GCPRECECS-----AOD 46
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 51 TTYDCNDLFLTANPPALPACTOILLQSNSTYVVDSEGLYANLVELDSONSFSDARD 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 47 RALVCHRRKRPVAVPEGIPETRIIDLGKNIKIKINODEFASFPHELEENINIVASVPR 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 CDFH-----ALPOLLSLHLENOLTRLEDHSHFALASLOEL 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 GAFNNIFNLRTGLBRNRLKILPGVFTGISNLTIKDISNKIVILLDYFODLYNKLSTL 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 147 YLNHNQLYIAPRAFSGLSNLRLHLNSMLRAIDSRWEMLPNLEIMIGKNKYDAIID 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 EVGDNDLYVISHRAFSGLNSLEQLEKCLGTSIPREALSHLGLVLRRLHNTAINARD 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 MNPRLANRSLVLGAM-NLRISDYALBGLSLESLSFYDNOANVPRLAEQVGLKF 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 YSFKRLYRLKVLKLEISHWPYLDITWTPNCLVGL-NLTSLSTHCHMTAVPYLAVERHLVYLRP 285
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 266 LDLNKRPDLQVGGDFANMLHLKEGLNNNEELVSDKALVNLPELTIKDITNNRSLF 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 LNSTYPISTIE-----GSMH-----ELRLQETQLV-----GGOLAM 319
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 326 IHPRAFHLPQMETMLNNALSALHQOVESLPNLQEVGLHGNPIRCDVCVIRMANATGT 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 VEPYAFARGINYLRLVNLVSGNQLTTEESVFHSGNLETLIDSNPLACCCORLIMVRRRR 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 RYRFTIEPOSTLCAEPDLDRLRYREVPFREMTHCLPLISPSNF-----PPSLQ 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 RLNFNRQOFT-CATPEFVGGKEKDFP-----DVLIPNYFTCRARIRDRKAQOYF 429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 435 VASGESMVHCRALAEPEEITYWVTPAGLRTPAHAGRYRYVPEGTLELRTATAEAGL 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 430 VDEGHTVQVCGAADDPPAIIIMLSR-KHLVSAKSNGLTYFPDSTLEKRAAOVDNGT 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 495 YTCVANTLVGADT-----KTVSVVGRALLQGRDEGQLELRVOE 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 YLCIANANGNDSPMHLHVRSYSPDMPHQPKMTFAFIN-----QPG--EGEANGSTRATV 542
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 536 TTPYHI 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 543 PPFEDI 548
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RA Magatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.,
RA "NEO" human cDNA sequencing project."
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL353746; CAC22713.1;
DR EMBL: AK056372; BAB71167.1;
DR InterPro: IPR003599; Ig_
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF00560; LRR_11.
DR Pfam: PF01463; LRRCT; 1.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PR00019; LEURICHNPT.
DR SMART: SM00409; Ig_1.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 10.
KW Immunoglobulin domain.
SQ
SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;

Query Match 12.6%; Score 468.5; DB 4; Length 606;
Best local similarity 27.0%; Pred. No. 2,7e-27;
Matches 140; Conservative 84; Mismatches 221; Indels 73; Gaps 9;

QY 27 VCPPOCAQOIRPWTPRSSYREATTVDGNDLFLAVPALPAGTQTLILQSNSTYRVQD 86
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 ICGPACCECS-----AQNKSVSCHRRRLIAIPGIPLETETK----- 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 SELGYANTLELDLQSNSTSDARDCDFHALPQLLSLHLEENQTRLEDHSHFALASLOEL 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 -----LDLSKNRLKSVNPEEFISYPLLEETLSDNTIANYEPAGNLFNLRLSL 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 147 YLNHNQLYIAPRAFSGLSNLRLHLNSMLRAIDSRWEMLPNLEIMIGKNKYDAIID 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 RLKGNRLKLVPGVFTGLSNLTIKDISNKIVILLDYFODLYNKLSEVGDNDLYISH 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 MNPRLANRSLVLGAM-NLRISDYALBGLSLESLSFYDNOANVPRLAEQVGLKF 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 RAFSGILSLEQLTLEKCNLTAVPTALSHLRSLISLHKLNINMPPYAFKRLFKHL 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 DLNKNPDLQVGGDFANMLHLKEGLNNNEELVSDKALVNLPELTIKDITNNP----- 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 EIDYMPLDMPANSLYGNLTSLSVTN-TNLSTVPELAFKHLVYTHLNTLSNPSTIE 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 322 -----RLSFTHPRAFHLPQMETMLNNALSALHQOVESLPNLQOE 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 290 AGMFSDLRLQELHIVGQALRTIEPHSFGGLFELVLANSONMLTELENNVSSPRLLEV 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 364 VCLHGNPIRCDVCVIRMANATGTRVPIEPQSTLCAEPDLDRLPYREVPFREMTHCLPL 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 ISINNPLACDCRLMIITLQROPTLOF-GGQOPMCAGPD-----TRESFKDFHSTALS 403
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 424 I-----SPRSFPSTLQ-----VASGESMVHCRALAEPEEITYWVTPAGLRTPAHAGRYRY 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 404 YFTCKKPKIREKKLOHLVDSQTVQLCSAGDPQVPSWTPRRRTTTSNGRA-TV 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 477 YPEGTLELRVYAEAGLYTCAONLVGADRTVSVVV 514
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 463 LGDTLEIRFQDDDSGMVVCIASAANDPTTASTLV 500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 14
P70193 PRELIMINARY; PRT; 1091 AA.

```


Query Match	11.7%	Score 438.5;	DB 4;	Length 640;
Best Local Similarity	25.5%;	Pred. No. 5.8e-25;		
Matches 148;	Conservative 84;	Mismatches 231;	Indels 117;	Gaps 12

```

0Y 2 RLTVAPLLLWAGATATVPVPMVHPVPCPOCAOIWRWTPRASSYREATVYDONDLEFLT 61
   | | | | |
Db 20 RALRFPRLVLLAIDLVLVAGLVARQICPSVSCSS-----NQSFKYICVRKNLR 68
   | | | | |
0Y 62 AVPPALPAGTOTLLLOSNIYVDOSELGYLANCTELDLSONSFSARDCCDFHALPOLLS 121
   | | | | |
Db 69 EYPPDQISINTRLNLHENQOIQIKVNSPKHLRIEILLQSNHIRTIEIGAFNLANT 128
   | | | | |
0Y 122 LHLBNOLTRLEDHSPAGLASLOELYLNHOLYIAPRATSGSLNLRHLNLSNLRAID 181
   | | | | |
Db 129 LELDNRLTPTTPNGAFVYLSKLELWLNMPNIESIPSAFNRIPSLRRDLGE----- 181
   | | | | |
0Y 182 SRWFEMLEPTEILMIGCKKVDAILDMNRPANRSLVLIAGMNRREISDVALBESLOS 241
   | | | | |
Db 182 -----LKRSTYISEGA-----FEGSLNRYLNLAMCNLRELPN----- 214
   | | | | |
0Y 242 LSFYDNOIARVBRRALEQVPLKFLDLINKNPLQVBGDFANMLHKELGJNNWELYSI 301
   | | | | |
Db 215 -----LPLIKLDELDLSGNLSIIRGSGOGLMHLQKMLMISO----- 254
   | | | | |
0Y 302 DKFALVNLPELTKLIDITNNPPLSTIHPRAFHILPOMETLIMNNALSALHQOVESLPNL 361
   | | | | |
Db 255 -----IQYIERNAFENQLOSLVEINLNAHNHNLTLPHDLETPPLHL 293
   | | | | |
0Y 362 QEVGHGNPIRCDCVIRMANATGRVAFIEQSTLCA-----EPDDLORLVREYPREMT 417
   | | | | |
Db 294 ERHJHNNHPMNCDDIIMLS--TWIDMAPSNPACCARCNTPNLKGRIYIGELDONYPT 350
   | | | | |
0Y 418 DHCEPLLSRSPESLOVAGSESWLRCRALAPEPEITVWTAGLITPRAHGRRRY 477
   | | | | |
Db 351 CYAPIYER---PADLWNTBGMALCKRA-STSLTYSVNTTPNGYVMTGATKYRIAYL 406
   | | | | |
0Y 478 PEGTLELRVTAEBAGLYTCAQNLVGADTKTVSVVVGRA-----LLOGR 523
   | | | | |
Db 407 SDGGLNFNFNTVQDTGMTCVNSVGNTPASATLNTATTTPEFSYFVETNEPQ 466
   | | | | |
0Y 524 DEGOGLELRVOETHRYHLLSWVPRTVSTNLWSSASS 563
   | | | | |
Db 467 DEARTDNNNGPTP---VYDMET--TNVSTTLTPOSTRS 500
   | | | | |

```

Search completed: August 31, 2002, 14:40:33
Job time: 260 sec